On Constrained Mixed-Integer DR-Submodular Minimization

Qimeng (Kim) Yu and Simge Küçükyavuz

Department of Industrial Engineering and Management Sciences
Northwestern University

Jan 12, 2023

Supported by ONR Grant N000142212602.

Classical Submodularity

$N = \{1, 2, \ldots, n\}$

Classical Submodularity

$N = \{1, 2, \ldots, n\}$

Definition (submodular set function)

A function $f : 2^N \to \mathbb{R}$ is **submodular** if $f(X) + f(Y) \geq f(X \cup Y) + f(X \cap Y)$ for any $X, Y \subseteq N$.

Classical Submodularity

$N = \{1, 2, \ldots, n\}$

Definition (submodular set function)

A function $f : 2^N \to \mathbb{R}$ is **submodular** if $f(X) + f(Y) \geq f(X \cup Y) + f(X \cap Y)$ for any $X, Y \subseteq N$.

Intuition. Submodularity $\Leftrightarrow$ Diminishing (Marginal) Returns (DR).

Classical Submodularity

Marginal return $\rho_i(S) := f(S \cup \{i\}) - f(S)$.

Classical Submodularity

Marginal return $\rho_i(S) := f(S \cup \{i\}) - f(S)$.

Definition (submodular set function, alternative)

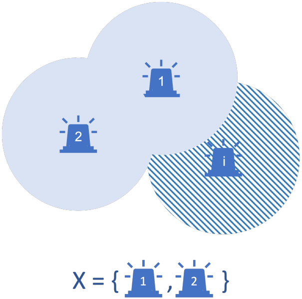
A function $f : 2^N \to \mathbb{R}$ is **submodular** if $\rho_i(X) \geq \rho_i(Y)$ for any $X \subseteq Y \subseteq N$, $i \in N \backslash Y$.

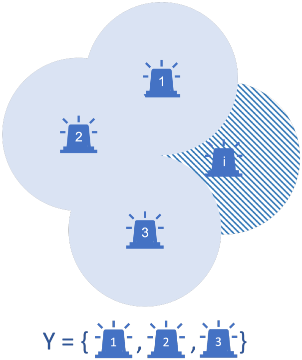
Classical Submodularity

Marginal return $\rho_i(S) := f(S \cup \{i\}) - f(S)$.

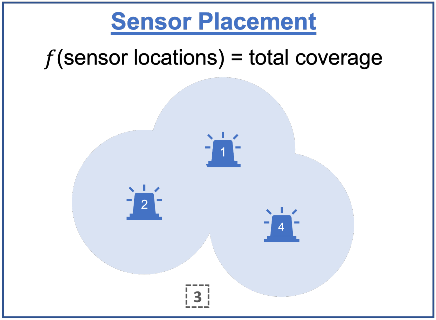
Definition (submodular set function, alternative)

A function $f : 2^N \to \mathbb{R}$ is **submodular** if $\rho_i(X) \geq \rho_i(Y)$ for any $X \subseteq Y \subseteq N$, $i \in N \setminus Y$.





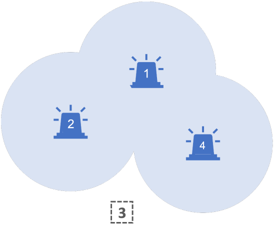
Classical Submodular Set Function Optimization



Classical Submodular Set Function Optimization

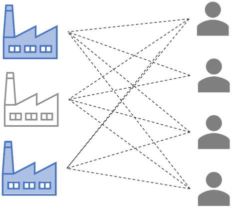
Sensor Placement

f(sensor locations) = total coverage



Facility Location

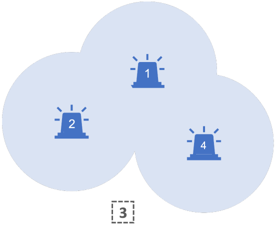
f(facility locations) = total benefit



Classical Submodular Set Function Optimization

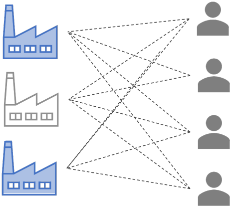
Sensor Placement

f(sensor locations) = total coverage



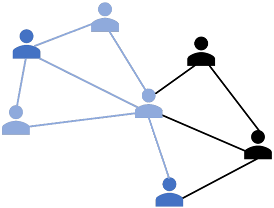
Facility Location

f(facility locations) = total benefit

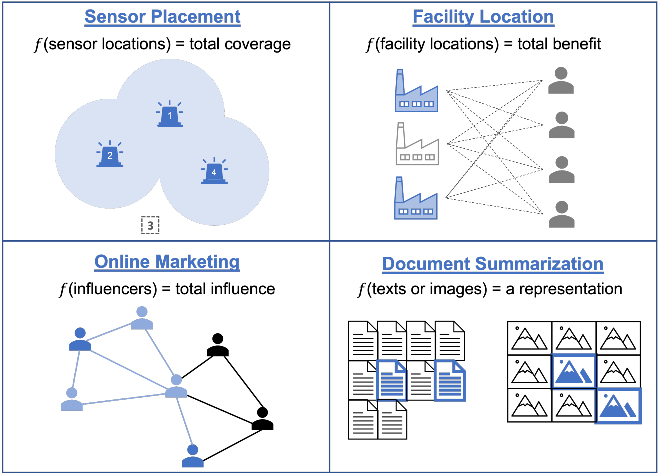


Online Marketing

f(influencers) = total influence



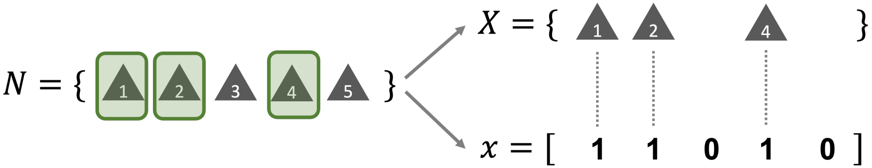
Classical Submodular Set Function Optimization



More applications: assortment optimization, mean-risk optimization, problems involving clustering, coverage, risk aversion, economies of scale, etc..

Classical Submodular Set Function Optimization

Decision space: Selection from a **single** ground set. Modeled by **binary** variables.



Slight abuse of notation: Use $f(X)$ and $f(x)$ interchangeably.

Submodular Minimization

Unconstrained Submodular Set Function Minimization

- Strongly poly-time solvable [e.g., Iwata et al., 2001, Orlin, 2009]
- conv (epigraph of f) is given by extended polymatroid inequalities (EPI) [Atamtürk and Narayanan, 2021, Lovász, 1983]
- Separation of EPIs is easy (greedy does it) [Edmonds, 1970]

Submodular Minimization

Unconstrained Submodular Set Function Minimization

- Strongly poly-time solvable [e.g., Iwata et al., 2001, Orlin, 2009]
- conv (epigraph of f) is given by extended polymatroid inequalities (EPI) [Atamtürk and Narayanan, 2021, Lovász, 1983]
- Separation of EPIs is easy (greedy does it) [Edmonds, 1970] $\rightarrow$ An equivalent (exponential) LP

Cardinality-constrained Submodular Set Function Minimization

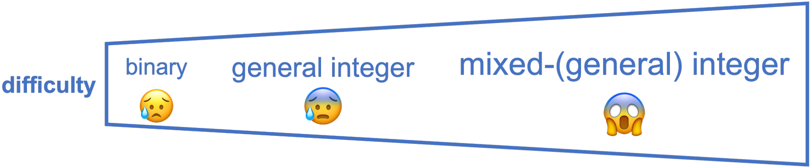
- NP-hard and hard to approximate (polynomial factor lower bounds on the approximation factor) [Svitkina and Fleischer, 2011]
- Use the EPIs to solve the resulting problem as an MILP (delayed cut generation) [e.g., Atamtürk and Narayanan, 2008, for mean-risk optimization]

Research question:

Exact approaches for optimizing generalizations of submodular set functions to **mixed-integer** variables (i.e., choose **multiple (discrete or continuous)** copies of each item)?

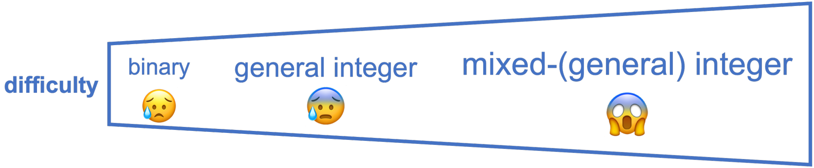
Research question:

Exact approaches for optimizing generalizations of submodular set functions to **mixed-integer** variables (i.e., choose **multiple (discrete or continuous)** copies of each item)?



Research question:

Exact approaches for optimizing generalizations of submodular set functions to **mixed-integer** variables (i.e., choose **multiple (discrete or continuous)** copies of each item)?



More formally,

- finitely convergent convexification schemes for mixed-binary or pure integer programs become infinitely convergent for mixed-integer programs.
- lifting problem of a valid inequality with mixed-binary variables is linear, whereas it is nonlinear for mixed-integer variables.

Diminishing Returns (DR)-Submodular Minimization

Diminishing Returns (DR)-Submodular Functions

$\mathbf{e}^i$: a vector with 1 in the ith entry, 0 elsewhere.

Definition (DR-submodular)

A function $f : \mathcal{X} \subseteq \mathbb{Z}^n \times \mathbb{R}^m \to \mathbb{R}$ is **DR-submodular** if

$$f(\mathbf{x} + \alpha \mathbf{e}^i) - f(\mathbf{x}) \geq f(\mathbf{y} + \alpha \mathbf{e}^i) - f(\mathbf{y})$$

for every $i \in \{1, 2, \ldots, n+m\}$, for all $\mathbf{x}, \mathbf{y} \in \mathcal{X}$ with $\mathbf{x} \leq \mathbf{y}$ component-wise, and for all $\alpha \in \mathbb{R}_+$ such that $\mathbf{x} + \alpha \mathbf{e}^i, \mathbf{y} + \alpha \mathbf{e}^i \in \mathcal{X}$.

Diminishing Returns (DR)-Submodular Functions

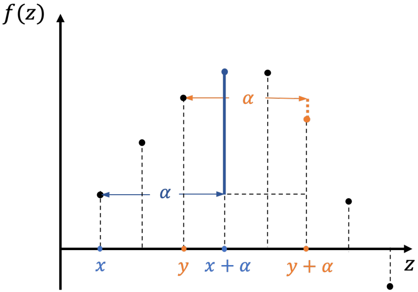
$\mathbf{e}^i$: a vector with 1 in the ith entry, 0 elsewhere.

Definition (DR-submodular)

A function $f : \mathcal{X} \subseteq \mathbb{Z}^n \times \mathbb{R}^m \to \mathbb{R}$ is **DR-submodular** if

$$f(\mathbf{x} + \alpha \mathbf{e}^i) - f(\mathbf{x}) \geq f(\mathbf{y} + \alpha \mathbf{e}^i) - f(\mathbf{y})$$

for every $i \in \{1, 2, \ldots, n+m\}$, for all $\mathbf{x}, \mathbf{y} \in \mathcal{X}$ with $\mathbf{x} \leq \mathbf{y}$ component-wise, and for all $\alpha \in \mathbb{R}_+$ such that $\mathbf{x} + \alpha \mathbf{e}^i, \mathbf{y} + \alpha \mathbf{e}^i \in \mathcal{X}$.



Diminishing Returns (DR)-Submodular Functions

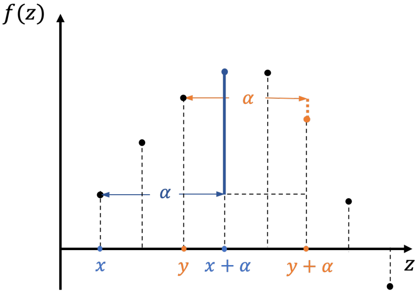
$\mathbf{e}^i$: a vector with 1 in the ith entry, 0 elsewhere.

Definition (DR-submodular)

A function $f : \mathcal{X} \subseteq \mathbb{Z}^n \times \mathbb{R}^m \to \mathbb{R}$ is **DR-submodular** if

$$f(\mathbf{x} + \alpha \mathbf{e}^i) - f(\mathbf{x}) \geq f(\mathbf{y} + \alpha \mathbf{e}^i) - f(\mathbf{y})$$

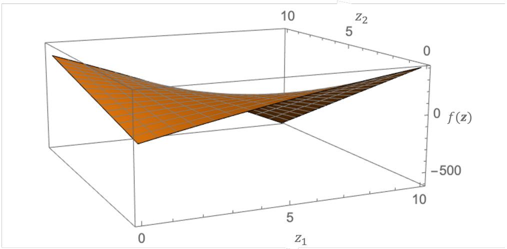
for every $i \in \{1, 2, \ldots, n+m\}$, for all $\mathbf{x}, \mathbf{y} \in \mathcal{X}$ with $\mathbf{x} \leq \mathbf{y}$ component-wise, and for all $\alpha \in \mathbb{R}_+$ such that $\mathbf{x} + \alpha \mathbf{e}^i, \mathbf{y} + \alpha \mathbf{e}^i \in \mathcal{X}$.



Note: DR-submodular functions are not necessarily concave.

DR-Submodular Functions

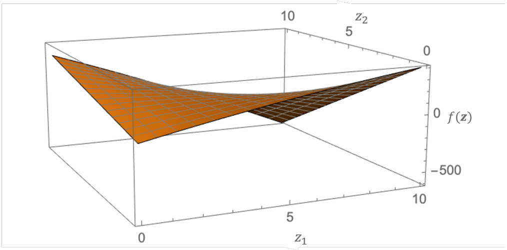
Example. Quadratic functions with non-positive Hessian entries (can be non-convex and non-concave).



A continuous DR-submodular function $f(\mathbf{z}) = -z_1^2 - 13z_1z_2 + 50z_1 + 30z_2$.

DR-Submodular Functions

Example. Quadratic functions with non-positive Hessian entries (can be non-convex and non-concave).

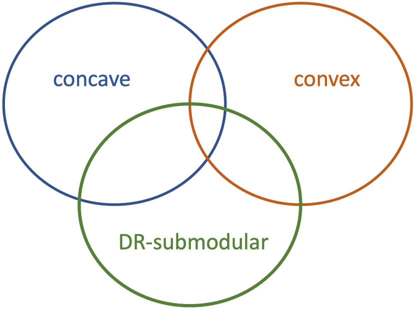


A continuous DR-submodular function $f(\mathbf{z}) = -z_1^2 - 13z_1z_2 + 50z_1 + 30z_2$.

Example. Submodular set functions when $\mathcal{X} = \{0, 1\}^n$.

Mixed-Integer DR-Submodular Optimization

Challenging!



Existing Literature

DR-submodular maximization

$$\max_{\mathbf{z} \in \mathcal{Z}} f(\mathbf{z})$$

- Pure integer: Ene and Nguyen [2016], Soma and Yoshida [2017], Soma and Yoshida [2018].
- Continuous: Bian et al. [2017], Ene and Nguyen [2020], Medal and Ahanor [2022], Niazadeh et al. [2018], Sadeghi and Fazel [2020].

Existing Literature

DR-submodular maximization

$$\max_{\mathbf{z}\in\mathcal{Z}} f(\mathbf{z})$$

- Pure integer: Ene and Nguyen [2016], Soma and Yoshida [2017], Soma and Yoshida [2018].
- Continuous: Bian et al. [2017], Ene and Nguyen [2020], Medal and Ahanor [2022], Niazadeh et al. [2018], Sadeghi and Fazel [2020].

DR-submodular minimization

$$\min_{\mathbf{z}\in\mathcal{Z}} f(\mathbf{z})$$

- Bach [2019], Ene and Nguyen [2016].
- Pseudo-polynomial algorithms, pure integer variables.

Existing Literature

DR-submodular maximization

$$\max_{\mathbf{z} \in \mathcal{Z}} f(\mathbf{z})$$

- Pure integer: Ene and Nguyen [2016], Soma and Yoshida [2017], Soma and Yoshida [2018].
- Continuous: Bian et al. [2017], Ene and Nguyen [2020], Medal and Ahanor [2022], Niazadeh et al. [2018], Sadeghi and Fazel [2020].

DR-submodular minimization

$$\min_{\mathbf{z} \in \mathcal{Z}} f(\mathbf{z})$$

- Bach [2019], Ene and Nguyen [2016].
- Pseudo-polynomial algorithms, pure integer variables.
- Questions: Mixed-integer variables?

Existing Literature

DR-submodular maximization

$$\max_{\mathbf{z} \in \mathcal{Z}} f(\mathbf{z})$$

- Pure integer: Ene and Nguyen [2016], Soma and Yoshida [2017], Soma and Yoshida [2018].
- Continuous: Bian et al. [2017], Ene and Nguyen [2020], Medal and Ahanor [2022], Niazadeh et al. [2018], Sadeghi and Fazel [2020].

DR-submodular minimization

$$\min_{\mathbf{z} \in \mathcal{Z}} f(\mathbf{z})$$

- Bach [2019], Ene and Nguyen [2016].
- Pseudo-polynomial algorithms, pure integer variables.
- Questions: Mixed-integer variables? Beyond box constraints?

Existing Literature

DR-submodular maximization

$$\max_{\mathbf{z}\in\mathcal{Z}} f(\mathbf{z})$$

- Pure integer: Ene and Nguyen [2016], Soma and Yoshida [2017], Soma and Yoshida [2018].
- Continuous: Bian et al. [2017], Ene and Nguyen [2020], Medal and Ahanor [2022], Niazadeh et al. [2018], Sadeghi and Fazel [2020].

DR-submodular minimization

$$\min_{\mathbf{z}\in\mathcal{Z}} f(\mathbf{z})$$

- Bach [2019], Ene and Nguyen [2016].
- Pseudo-polynomial algorithms, pure integer variables.
- Questions: Mixed-integer variables? Beyond box constraints? Polynomial-time solvable?

Problem Description

$$\min_{\mathbf{z}\in\mathcal{Z}(\mathcal{G},\mathbf{u})} f(\mathbf{z}),$$

where f is DR-submodular and $\mathcal{G} = (\mathcal{V}, \mathcal{A})$ is a DAG representing the monotonicity relations among variables in $\mathcal{V} = \{1, 2, \ldots, n+m\}$.

Problem Description

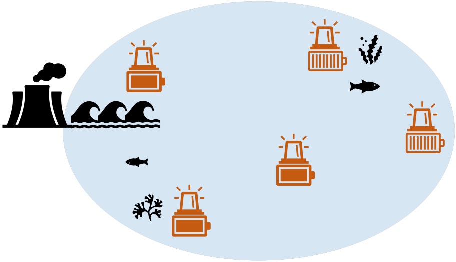
$$\min_{\mathbf{z}\in\mathcal{Z}(\mathcal{G},\mathbf{u})} f(\mathbf{z}),$$

where f is DR-submodular and $\mathcal{G}=(\mathcal{V},\mathcal{A})$ is a DAG representing the monotonicity relations among variables in $\mathcal{V}=\{1,2,\ldots,n+m\}$.

$$\mathcal{Z}(\mathcal{G},\mathbf{u}) := \{\mathbf{z}\in\mathbb{Z}^n\times\mathbb{R}^m : \mathbf{0}\leq\mathbf{z}\leq\mathbf{u}, z_i\leq z_j,\ \forall\,(i,j)\in\mathcal{A}\}.$$

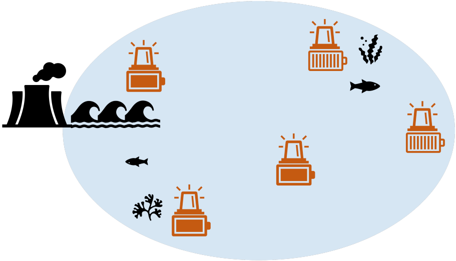
Mixed-Integer DR-Submodular Example

Sensor energy management.



Mixed-Integer DR-Submodular Example

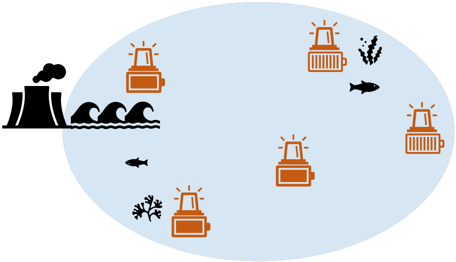
Sensor energy management.



- z_i: energy level of sensor at location $i \in \mathcal{V}$ (**continuous or discrete**)

Mixed-Integer DR-Submodular Example

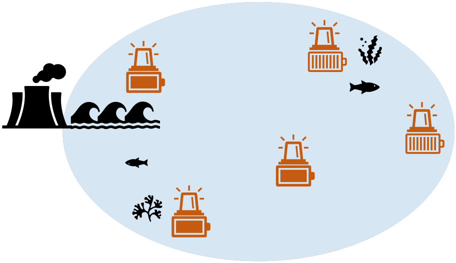
Sensor energy management.



- z_i: energy level of sensor at location $i \in \mathcal{V}$ (**continuous or discrete**)
- $f(\mathbf{z}) = \mathbb{E}$[saved detection time] is **DR-submodular**

Mixed-Integer DR-Submodular Example

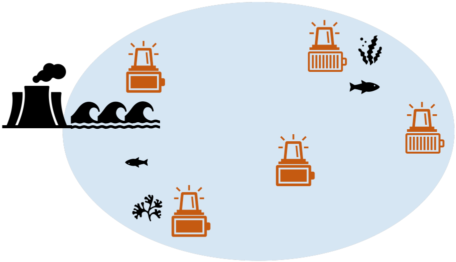
Sensor energy management.



- z_i: energy level of sensor at location $i \in \mathcal{V}$ (**continuous or discrete**)
- $f(\mathbf{z}) = \mathbb{E}[\text{saved detection time}]$ is **DR-submodular**
  $$= t_\infty - \sum_{\omega \in \Omega} p_\omega \sum_{S \subseteq \mathcal{V}} (\min_{i \in S} t_{\omega,i}) \prod_{i \in S} (1 - (1-p)^{z_i}) \prod_{i \notin S} (1-p)^{z_i}$$
  - Ω: contamination events; $\omega \in \Omega$ has probability p_ω
  - $1 - (1-p)^{z_i}$: chance of successful detection of any event by i at energy level z_i
  - $t_{\omega,i}$: time to detect event ω by sensor i; $t_\infty := \max_{i \in \mathcal{V}, \omega \in \Omega} t_{\omega,i}$.

Mixed-Integer DR-Submodular Example

Sensor energy management.



- z_i: energy level of sensor at location $i \in \mathcal{V}$ (**continuous or discrete**)
- $f(\mathbf{z}) = \mathbb{E}[\text{saved detection time}]$ is **DR-submodular**
  $$= t_\infty - \sum_{\omega \in \Omega} p_\omega \sum_{S \subseteq \mathcal{V}} (\min_{i \in S} t_{\omega,i}) \prod_{i \in S} (1 - (1-p)^{z_i}) \prod_{i \notin S} (1-p)^{z_i}$$
  - Ω: contamination events; $\omega \in \Omega$ has probability p_ω
  - $1 - (1-p)^{z_i}$: chance of successful detection of any event by i at energy level z_i
  - $t_{\omega,i}$: time to detect event ω by sensor i; $t_\infty := \max_{i \in \mathcal{V}, \omega \in \Omega} t_{\omega,i}$.
- Some sensor information is of higher priority $\Rightarrow$ **monotonicity constraints**.

Problem Description

$$\mathcal{Z}(\mathcal{G}, \mathbf{u}) := \{\mathbf{z} \in \mathbb{Z}^n \times \mathbb{R}^m : \mathbf{0} \leq \mathbf{z} \leq \mathbf{u}, z_i \leq z_j, \ \forall\, (i,j) \in \mathcal{A}\}.$$

$\mathcal{G} = (\mathcal{V}, \mathcal{A})$: a **directed rooted forest**.

Problem Description

$$\mathcal{Z}(\mathcal{G}, \mathbf{u}) := \{\mathbf{z} \in \mathbb{Z}^n \times \mathbb{R}^m : \mathbf{0} \leq \mathbf{z} \leq \mathbf{u}, z_i \leq z_j, \ \forall (i,j) \in \mathcal{A}\}.$$

$\mathcal{G} = (\mathcal{V}, \mathcal{A})$: a **directed rooted forest**.

Definition (Directed rooted forest)

A disjoint union of directed rooted tree(s) with arcs pointing away from the root(s).

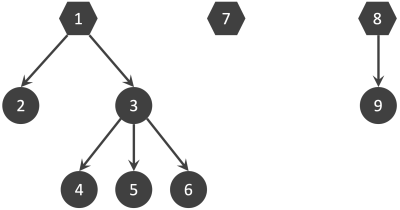
Problem Description

$$\mathcal{Z}(\mathcal{G}, \mathbf{u}) := \{\mathbf{z} \in \mathbb{Z}^n \times \mathbb{R}^m : \mathbf{0} \leq \mathbf{z} \leq \mathbf{u}, z_i \leq z_j, \ \forall\, (i,j) \in \mathcal{A}\}.$$

$\mathcal{G} = (\mathcal{V}, \mathcal{A})$: a **directed rooted forest**.

Definition (Directed rooted forest)

A disjoint union of directed rooted tree(s) with arcs pointing away from the root(s).



Example A

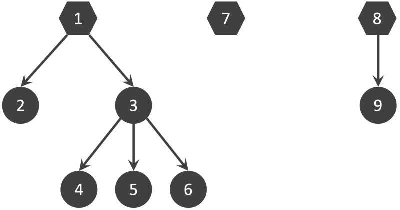
Problem Description

$$\mathcal{Z}(\mathcal{G}, \mathbf{u}) := \{\mathbf{z} \in \mathbb{Z}^n \times \mathbb{R}^m : \mathbf{0} \leq \mathbf{z} \leq \mathbf{u}, z_i \leq z_j, \forall (i,j) \in \mathcal{A}\}.$$

$\mathcal{G} = (\mathcal{V}, \mathcal{A})$: a **directed rooted forest**.

Definition (Directed rooted forest)

A disjoint union of directed rooted tree(s) with arcs pointing away from the root(s).



Example A

Example B

Equivalent Formulation

Recall

$$\min_{\mathbf{z}\in\mathcal{Z}(\mathcal{G},\mathbf{u})} f(\mathbf{z}),$$

$$\mathcal{Z}(\mathcal{G},\mathbf{u}) := \{\mathbf{z}\in\mathbb{Z}^n\times\mathbb{R}^m : \mathbf{0}\le\mathbf{z}\le\mathbf{u}, z_i\le z_j,\ \forall\,(i,j)\in\mathcal{A}\}.$$

Equivalent Formulation

Recall

$$\min_{\mathbf{z}\in\mathcal{Z}(\mathcal{G},\mathbf{u})} f(\mathbf{z}),$$

$$\mathcal{Z}(\mathcal{G},\mathbf{u}) := \{\mathbf{z} \in \mathbb{Z}^n \times \mathbb{R}^m : \mathbf{0} \le \mathbf{z} \le \mathbf{u}, z_i \le z_j, \ \forall\, (i,j) \in \mathcal{A}\}.$$

Equivalently,

$$\min \left\{ w : (\mathbf{z}, w) \in \operatorname{conv}\left(\mathcal{P}_f^{\mathcal{Z}(\mathcal{G},\mathbf{u})}\right) \right\},$$

where

$$\mathcal{P}_f^{\mathcal{Z}(\mathcal{G},\mathbf{u})} := \{(\mathbf{z}, w) \in \mathcal{Z}(\mathcal{G},\mathbf{u}) \times \mathbb{R} : w \ge f(\mathbf{z})\}.$$

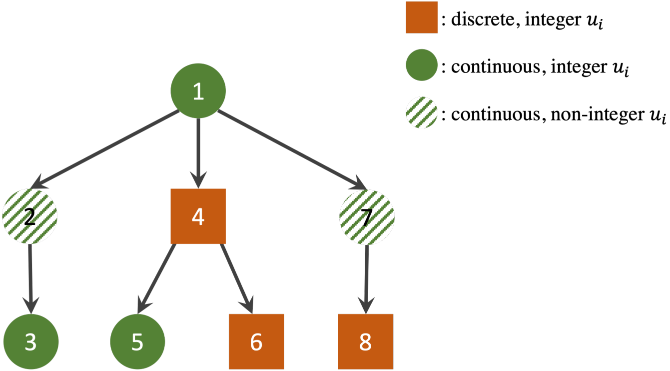
Understand $\operatorname{conv}(\mathcal{Z}(\mathcal{G}, \mathbf{u}))$

Observation. Continuous relaxation of $\mathcal{Z}(\mathcal{G}, \mathbf{u})$ is **not** necessarily $\operatorname{conv}(\mathcal{Z}(\mathcal{G}, \mathbf{u}))$.

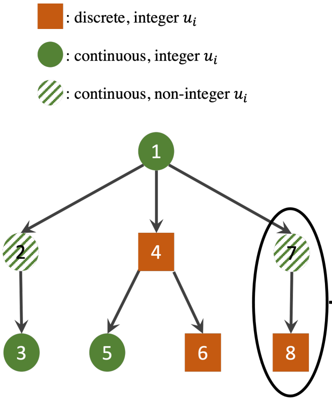
Understand conv $(\mathcal{Z}(\mathcal{G}, \mathbf{u}))$

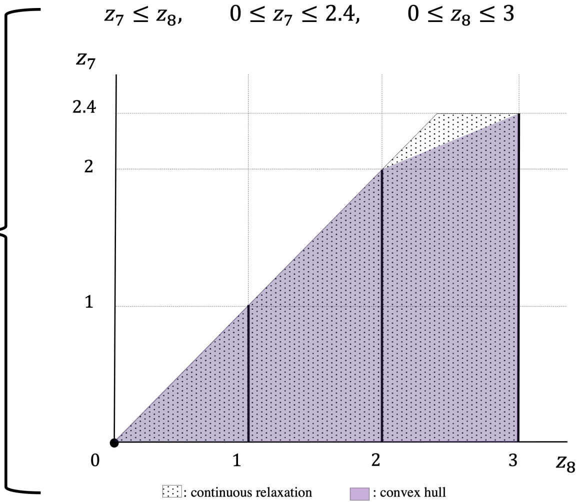
Observation. Continuous relaxation of $\mathcal{Z}(\mathcal{G}, \mathbf{u})$ is **not** necessarily conv $(\mathcal{Z}(\mathcal{G}, \mathbf{u}))$.

Example.



Understand conv $(\mathcal{Z}(\mathcal{G}, \mathbf{u}))$

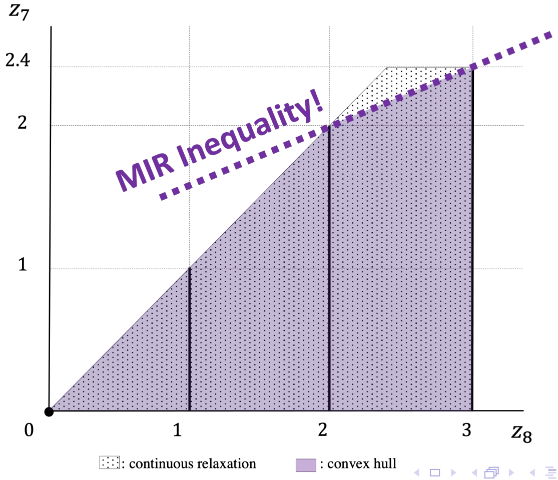




Understand conv $(\mathcal{Z}(\mathcal{G}, \mathbf{u}))$

Mixed-Integer Rounding (MIR) inequality [Nemhauser and Wolsey, 1990]:

$$-z_8 + \frac{z_7}{u_7 - \lfloor u_7 \rfloor} \leq \frac{\lfloor u_7 \rfloor (\lceil u_7 \rceil - u_7)}{u_7 - \lfloor u_7 \rfloor}.$$



Understand $\operatorname{conv}(\mathcal{Z}(\mathcal{G}, \mathbf{u}))$

$\Psi =$ The set of fractionally upper-bounded continuous variables with discrete descendant(s).

Theorem (informal; full description of $\operatorname{conv}(\mathcal{Z}(\mathcal{G}, \mathbf{u}))$) [Yu and Küçükyavuz, 2022]

Under some conditions, $\operatorname{conv}(\mathcal{Z}(\mathcal{G}, \mathbf{u}))$ is fully described by the trivial inequalities and the MIR inequalities for all $\psi \in \Psi$ and their children:

$$-z_{\mathrm{ch}(\psi)} + \frac{z_\psi}{u_\psi - \lfloor u_\psi \rfloor} \leq \frac{\lfloor u_\psi \rfloor (\lceil u_\psi \rceil - u_\psi)}{u_\psi - \lfloor u_\psi \rfloor}.$$

Characterize conv $\left(\mathcal{P}_f^{\mathcal{Z}(\mathcal{G},\mathbf{u})}\right)$

Proposition (validity of DR inequalities) [Yu and Küçükyavuz, 2022]

For *certain* permutations $\delta = (\delta(1), \delta(2), \dots, \delta(|\mathcal{V}|))$, a **DR inequality associated with** δ

$$w \geq \sum_{k=0}^{|\mathcal{V}|} [t(\delta, \mathbf{z})_k - t(\delta, \mathbf{z})_{k+1}]\, f(P(\mathcal{T}^{\delta,k})).$$

is valid for $\mathcal{P}_f^{\mathcal{Z}(\mathcal{G},\mathbf{u})}$.

Characterize conv $\left(\mathcal{P}_f^{\mathcal{Z}(\mathcal{G},\mathbf{u})}\right)$

DR inequality [Yu and Küçükyavuz, 2022]

For *certain* permutation $\delta = (\delta(1), \delta(2), \ldots, \delta(|\mathcal{V}|))$, a **DR inequality associated with** δ is

$$w \geq \sum_{k=0}^{|\mathcal{V}|} [t(\delta, \mathbf{z})_k - t(\delta, \mathbf{z})_{k+1}]\, f(P(\mathcal{T}^{\delta,k})).$$

Characterize conv $\left(\mathcal{P}_f^{\mathcal{Z}(\mathcal{G},\mathbf{u})}\right)$

DR inequality [Yu and Küçükyavuz, 2022]

For *certain* permutation $\delta = (\delta(1), \delta(2), \dots, \delta(|\mathcal{V}|))$, a **DR inequality associated with** δ is

$$w \geq \sum_{k=0}^{|\mathcal{V}|} [t(\delta, \mathbf{z})_k - t(\delta, \mathbf{z})_{k+1}]\, f(P(\mathcal{T}^{\delta,k})).$$

Permutation

$\delta = (1,3,2)$

Characterize conv $\left(\mathcal{P}_f^{\mathcal{Z}(\mathcal{G},\mathbf{u})}\right)$

DR inequality [Yu and Küçükyavuz, 2022]

For *certain* permutation $\delta = (\delta(1), \delta(2), \ldots, \delta(|\mathcal{V}|))$, a **DR inequality associated with** δ is

$$w \geq \sum_{k=0}^{|\mathcal{V}|} [t(\delta, \mathbf{z})_k - t(\delta, \mathbf{z})_{k+1}] \, f(P(\mathcal{T}^{\delta,k})).$$

Permutation → Subsets of V

$\delta = (1,3,2)$

$T^{\delta,0} = \emptyset$

$T^{\delta,1} = \{1\}$

$T^{\delta,2} = \{1,3\}$

$T^{\delta,3} = \{1,2,3\}$

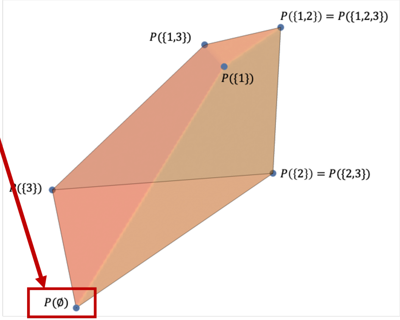
Characterize conv $\left(\mathcal{P}_f^{\mathcal{Z}(\mathcal{G},\mathbf{u})}\right)$

DR inequality [Yu and Küçükyavuz, 2022]

For *certain* permutation $\delta = (\delta(1), \delta(2), \ldots, \delta(|\mathcal{V}|))$, a **DR inequality associated with** δ is

$$w \geq \sum_{k=0}^{|\mathcal{V}|} [t(\delta, \mathbf{z})_k - t(\delta, \mathbf{z})_{k+1}]\, f(P(\mathcal{T}^{\delta,k})).$$

Permutation	Subsets of V	Extreme points of conv($\mathcal{Z}(\mathcal{G},\mathbf{u})$)
$\delta = (1,3,2)$	$\boxed{T^{\delta,0} = \emptyset}$	
	$T^{\delta,1} = \{1\}$	
	$T^{\delta,2} = \{1,3\}$	
	$T^{\delta,3} = \{1,2,3\}$	



Characterize conv $\left(\mathcal{P}_f^{\mathcal{Z}(\mathcal{G},\mathbf{u})}\right)$

DR inequality [Yu and Küçükyavuz, 2022]

For *certain* permutation $\delta = (\delta(1), \delta(2), \ldots, \delta(|\mathcal{V}|))$, a **DR inequality associated with** δ is

$$w \geq \sum_{k=0}^{|\mathcal{V}|} [t(\delta, \mathbf{z})_k - t(\delta, \mathbf{z})_{k+1}]\, f(P(\mathcal{T}^{\delta,k})).$$

Permutation → Subsets of V → Extreme points of conv($\mathcal{Z}(\mathcal{G},\mathbf{u})$)

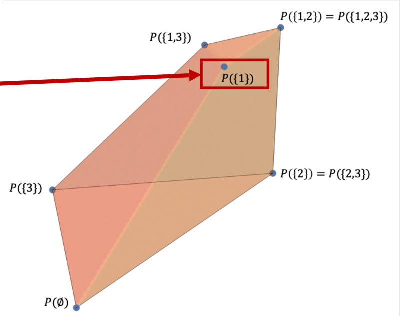
$\delta = (1,3,2)$

$T^{\delta,0} = \emptyset$

$\boxed{T^{\delta,1} = \{1\}}$

$T^{\delta,2} = \{1,3\}$

$T^{\delta,3} = \{1,2,3\}$



Characterize conv $\left(\mathcal{P}_f^{\mathcal{Z}(\mathcal{G},\mathbf{u})}\right)$

DR inequality [Yu and Küçükyavuz, 2022]

For *certain* permutation $\delta = (\delta(1), \delta(2), \ldots, \delta(|\mathcal{V}|))$, a **DR inequality associated with** δ is

$$w \geq \sum_{k=0}^{|\mathcal{V}|} [t(\delta, \mathbf{z})_k - t(\delta, \mathbf{z})_{k+1}]\, f(P(\mathcal{T}^{\delta,k})).$$

Permutation → Subsets of V → Extreme points of conv($\mathcal{Z}(\mathcal{G},\mathbf{u})$)

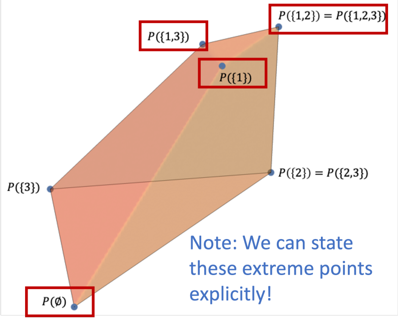
$\delta = (1,3,2)$

$\boxed{T^{\delta,0} = \emptyset}$

$\boxed{T^{\delta,1} = \{1\}}$

$\boxed{T^{\delta,2} = \{1,3\}}$

$\boxed{T^{\delta,3} = \{1,2,3\}}$



Note: We can state these extreme points explicitly!

Characterize conv $\left(\mathcal{P}_f^{\mathcal{Z}(\mathcal{G},\mathbf{u})}\right)$

DR inequality [Yu and Küçükyavuz, 2022]

For *certain* permutation $\delta = (\delta(1), \delta(2), \ldots, \delta(|\mathcal{V}|))$, a **DR inequality associated with** δ is

$$w \geq \sum_{k=0}^{|\mathcal{V}|} [t(\delta, \mathbf{z})_k - t(\delta, \mathbf{z})_{k+1}]\, f(P(\mathcal{T}^{\delta,k})).$$

Characterize conv $\left(\mathcal{P}_f^{\mathcal{Z}(\mathcal{G},\mathbf{u})}\right)$

DR inequality [Yu and Küçükyavuz, 2022]

For *certain* permutation $\delta = (\delta(1), \delta(2), \ldots, \delta(|\mathcal{V}|))$, a **DR inequality associated with** δ is

$$w \geq \sum_{k=0}^{|\mathcal{V}|} [t(\delta, \mathbf{z})_k - t(\delta, \mathbf{z})_{k+1}]\, f(P(\mathcal{T}^{\delta,k})).$$

$t(\delta, \mathbf{z})_k$: linear expression of $\mathbf{z}$, with explicit form.

Characterize conv $\left(\mathcal{P}_f^{\mathcal{Z}(\mathcal{G},\mathbf{u})}\right)$

DR inequality [Yu and Küçükyavuz, 2022]

For *certain* permutation $\delta = (\delta(1), \delta(2), \ldots, \delta(|\mathcal{V}|))$, a **DR inequality associated with** δ is

$$w \geq \sum_{k=0}^{|\mathcal{V}|} [t(\delta, \mathbf{z})_k - t(\delta, \mathbf{z})_{k+1}] \, f(P(\mathcal{T}^{\delta,k})).$$

$t(\delta, \mathbf{z})_k$: linear expression of $\mathbf{z}$, with explicit form.

DR inequalities are **linear** and **homogeneous**.

Characterize $\operatorname{conv}\left(\mathcal{P}_f^{\mathcal{Z}(\mathcal{G},\mathbf{u})}\right)$

DR inequality [Yu and Küçükyavuz, 2022]

For *certain* permutation $\delta = (\delta(1), \delta(2), \ldots, \delta(|\mathcal{V}|))$, a **DR inequality associated with** δ is

$$w \geq \sum_{k=0}^{|\mathcal{V}|} [t(\delta, \mathbf{z})_k - t(\delta, \mathbf{z})_{k+1}]\, f(P(\mathcal{T}^{\delta,k})).$$

$t(\delta, \mathbf{z})_k$: linear expression of $\mathbf{z}$, with explicit form.

DR inequalities are **linear** and **homogeneous**.

They subsume the well-known extended polymatroid inequalities [Atamtürk and Narayanan, 2021, Edmonds, 1970, Lovász, 1983].

Note the equivalent DR-inequality

$$w \geq \sum_{k=1}^{|\mathcal{V}|} t(\delta, \mathbf{z})_k [f(P(\mathcal{T}^{\delta,k})) - f(P(\mathcal{T}^{\delta,k-1}))].$$

An Important Property of $\operatorname{conv}(\mathcal{Z}(\mathcal{G}, \mathbf{u}))$

Proposition [Yu and Küçükyavuz, 2022]

For any $\bar{\mathbf{z}} \in \operatorname{conv}(\mathcal{Z}(\mathcal{G}, \mathbf{u}))$, let $\delta \leftarrow \texttt{Permutation\_Finder}(\bar{\mathbf{z}}, \mathcal{G}, \mathbf{u})$. Then $\bar{\mathbf{z}}$ can be written as the convex combination:

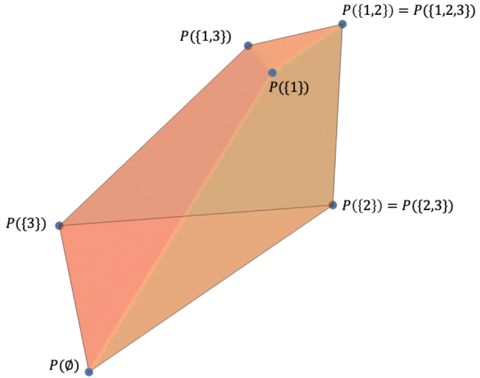
$$\bar{\mathbf{z}} = \sum_{k=0}^{|\mathcal{V}|} [t(\delta, \bar{\mathbf{z}})_k - t(\delta, \bar{\mathbf{z}})_{k+1}] \, P(\mathcal{T}^{\delta,k}).$$

An Important Property of conv$(\mathcal{Z}(\mathcal{G}, \mathbf{u}))$

Proposition [Yu and Küçükyavuz, 2022]

For any $\bar{\mathbf{z}} \in \text{conv}(\mathcal{Z}(\mathcal{G}, \mathbf{u}))$, let $\delta \leftarrow \texttt{Permutation\_Finder}(\bar{\mathbf{z}}, \mathcal{G}, \mathbf{u})$. Then $\bar{\mathbf{z}}$ can be written as the convex combination:

$$\bar{\mathbf{z}} = \sum_{k=0}^{|\mathcal{V}|} [t(\delta, \bar{\mathbf{z}})_k - t(\delta, \bar{\mathbf{z}})_{k+1}] \, P(\mathcal{T}^{\delta,k}).$$

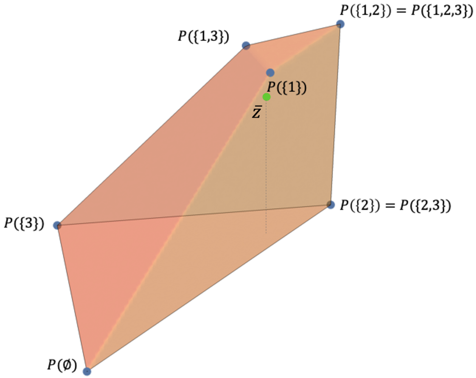


An Important Property of conv $(\mathcal{Z}(\mathcal{G}, \mathbf{u}))$

Proposition [Yu and Küçükyavuz, 2022]

For any $\bar{\mathbf{z}} \in \text{conv}\left(\mathcal{Z}(\mathcal{G}, \mathbf{u})\right)$, let $\delta \leftarrow \texttt{Permutation\_Finder}(\bar{\mathbf{z}}, \mathcal{G}, \mathbf{u})$. Then $\bar{\mathbf{z}}$ can be written as the convex combination:

$$\bar{\mathbf{z}} = \sum_{k=0}^{|\mathcal{V}|} [t(\delta, \bar{\mathbf{z}})_k - t(\delta, \bar{\mathbf{z}})_{k+1}] \, P(\mathcal{T}^{\delta,k}).$$



An Important Property of conv $(\mathcal{Z}(\mathcal{G}, \mathbf{u}))$

Proposition [Yu and Küçükyavuz, 2022]

For any $\bar{\mathbf{z}} \in \operatorname{conv}(\mathcal{Z}(\mathcal{G}, \mathbf{u}))$, let $\delta \leftarrow$ `Permutation_Finder`$(\bar{\mathbf{z}}, \mathcal{G}, \mathbf{u})$. Then $\bar{\mathbf{z}}$ can be written as the convex combination:

$$\bar{\mathbf{z}} = \sum_{k=0}^{|\mathcal{V}|} [t(\delta, \bar{\mathbf{z}})_k - t(\delta, \bar{\mathbf{z}})_{k+1}] \, P(\mathcal{T}^{\delta,k}).$$

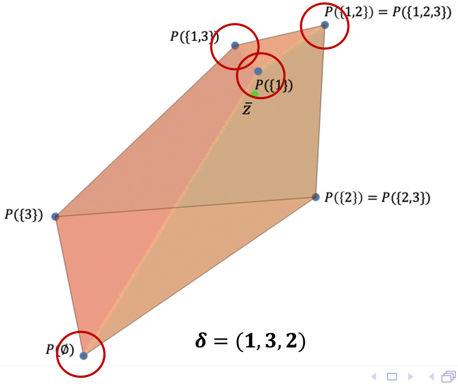
$$\bar{z} \longrightarrow \boxed{\texttt{Permutation\_Finder}} \longrightarrow \boldsymbol{\delta} = (\mathbf{1}, \mathbf{3}, \mathbf{2})$$

An Important Property of conv $(\mathcal{Z}(\mathcal{G}, \mathbf{u}))$

Proposition [Yu and Küçükyavuz, 2022]

For any $\bar{\mathbf{z}} \in \text{conv}\,(\mathcal{Z}(\mathcal{G}, \mathbf{u}))$, let $\delta \leftarrow \texttt{Permutation\_Finder}(\bar{\mathbf{z}}, \mathcal{G}, \mathbf{u})$. Then $\bar{\mathbf{z}}$ can be written as the convex combination:

$$\bar{\mathbf{z}} = \sum_{k=0}^{|\mathcal{V}|} [t(\delta, \bar{\mathbf{z}})_k - t(\delta, \bar{\mathbf{z}})_{k+1}]\, P(\mathcal{T}^{\delta,k}).$$

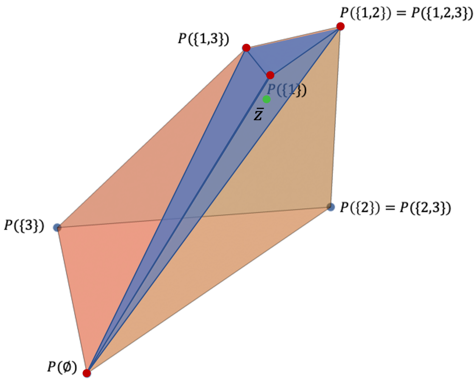


An Important Property of conv $(\mathcal{Z}(\mathcal{G}, \mathbf{u}))$

Proposition [Yu and Küçükyavuz, 2022]

For any $\bar{\mathbf{z}} \in \text{conv}\,(\mathcal{Z}(\mathcal{G}, \mathbf{u}))$, let $\delta \leftarrow \texttt{Permutation\_Finder}(\bar{\mathbf{z}}, \mathcal{G}, \mathbf{u})$. Then $\bar{\mathbf{z}}$ can be written as the convex combination:

$$\bar{\mathbf{z}} = \sum_{k=0}^{|\mathcal{V}|} [t(\delta, \bar{\mathbf{z}})_k - t(\delta, \bar{\mathbf{z}})_{k+1}]\, P(\mathcal{T}^{\delta,k}).$$



An Important Property of conv $(\mathcal{Z}(\mathcal{G}, \mathbf{u}))$

Proposition [Yu and Küçükyavuz, 2022]

For any $\bar{\mathbf{z}} \in \operatorname{conv}(\mathcal{Z}(\mathcal{G}, \mathbf{u}))$, let $\delta \leftarrow \texttt{Permutation\_Finder}(\bar{\mathbf{z}}, \mathcal{G}, \mathbf{u})$. Then $\bar{\mathbf{z}}$ can be written as the convex combination:

$$\bar{\mathbf{z}} = \sum_{k=0}^{|\mathcal{V}|} [t(\delta, \bar{\mathbf{z}})_k - t(\delta, \bar{\mathbf{z}})_{k+1}] \, P(\mathcal{T}^{\delta,k}).$$

For $k \in \{1, \ldots, |\mathcal{V}|\}$, let $i = \delta(k)$,

$$t(\delta, \mathbf{z})_k := \begin{cases} \dfrac{\eta_\psi(\mathbf{z}) - z_{i^\triangle(\mathcal{T}^{\delta,k-1})}}{\lfloor u_\psi \rfloor - u_{i^\triangle(\mathcal{T}^{\delta,k-1})}}, & \text{if } i = \psi \in \Psi \text{ and } \operatorname{ch}(\psi) \notin \mathcal{T}^{\delta,k-1}, \\ \dfrac{z_i - \eta_\psi(\mathbf{z})}{u_i - \lfloor u_\psi \rfloor}, & \text{else if } i^\triangle(\mathcal{T}^{\delta,k-1}) = \psi \in \Psi, \\ \dfrac{z_i - z_{i^\triangle(\mathcal{T}^{\delta,k-1})}}{u_i - u_{i^\triangle(\mathcal{T}^{\delta,k-1})}}, & \text{otherwise.} \end{cases}$$

$t(\delta, \mathbf{z})_0 = 1, \; t(\delta, \mathbf{z})_{|\mathcal{V}|+1} = 0.$

For any $\psi \in \Psi$, $\eta_\psi(\mathbf{z}) = \frac{z_\psi - (u_\psi - \lfloor u_\psi \rfloor) z_{\operatorname{ch}(\psi)}}{u_{\operatorname{ch}(\psi)} - u_\psi}$.

An Important Property of conv $(\mathcal{Z}(\mathcal{G}, \mathbf{u}))$

Proposition [Yu and Küçükyavuz, 2022]

For any $\bar{\mathbf{z}} \in \operatorname{conv}(\mathcal{Z}(\mathcal{G}, \mathbf{u}))$, let $\delta \leftarrow \texttt{Permutation\_Finder}(\bar{\mathbf{z}}, \mathcal{G}, \mathbf{u})$. Then $\bar{\mathbf{z}}$ can be written as the convex combination:

$$\bar{\mathbf{z}} = \sum_{k=0}^{|\mathcal{V}|} [t(\delta, \bar{\mathbf{z}})_k - t(\delta, \bar{\mathbf{z}})_{k+1}] \, P(\mathcal{T}^{\delta,k}).$$

For $k \in \{1, \ldots, |\mathcal{V}|\}$, let $i = \delta(k)$,

$$t(\delta, \mathbf{z})_k := \begin{cases} \dfrac{\eta_\psi(\mathbf{z}) - z_{i^\triangle(\mathcal{T}^{\delta,k-1})}}{\lfloor u_\psi \rfloor - u_{i^\triangle(\mathcal{T}^{\delta,k-1})}}, & \text{if } i = \psi \in \Psi \text{ and } \operatorname{ch}(\psi) \notin \mathcal{T}^{\delta,k-1}, \\ \dfrac{z_i - \eta_\psi(\mathbf{z})}{u_i - \lfloor u_\psi \rfloor}, & \text{else if } i^\triangle(\mathcal{T}^{\delta,k-1}) = \psi \in \Psi, \\ \dfrac{z_i - z_{i^\triangle(\mathcal{T}^{\delta,k-1})}}{u_i - u_{i^\triangle(\mathcal{T}^{\delta,k-1})}}, & \text{otherwise.} \end{cases}$$

$t(\delta, \mathbf{z})_0 = 1, \; t(\delta, \mathbf{z})_{|\mathcal{V}|+1} = 0.$

For any $\psi \in \Psi$, $\eta_\psi(\mathbf{z}) = \frac{z_\psi - (u_\psi - \lfloor u_\psi \rfloor) z_{\operatorname{ch}(\psi)}}{u_{\operatorname{ch}(\psi)} - u_\psi}$.

Gist: $t(\delta, \mathbf{z})_k$ is a linear expression of $\mathbf{z}$ that we can explicitly state.

Characterize $\operatorname{conv}\left(\mathcal{P}_f^{\mathcal{Z}(\mathcal{G},\mathbf{u})}\right)$

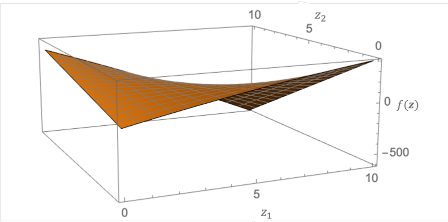
Theorem (full description of $\operatorname{conv}\left(\mathcal{P}_f^{\mathcal{Z}(\mathcal{G},\mathbf{u})}\right)$) [Yu and Küçükyavuz, 2022]

The DR inequalities, MIR inequalities, along with the box and monotonicity constraints, fully describe $\operatorname{conv}\left(\mathcal{P}_f^{\mathcal{Z}(\mathcal{G},\mathbf{u})}\right)$.

Characterize conv $\left(\mathcal{P}_f^{\mathcal{Z}(\mathcal{G},\mathbf{u})}\right)$

Theorem (full description of conv $\left(\mathcal{P}_f^{\mathcal{Z}(\mathcal{G},\mathbf{u})}\right)$) [Yu and Küçükyavuz, 2022]

The DR inequalities, MIR inequalities, along with the box and monotonicity constraints, fully describe conv $\left(\mathcal{P}_f^{\mathcal{Z}(\mathcal{G},\mathbf{u})}\right)$.

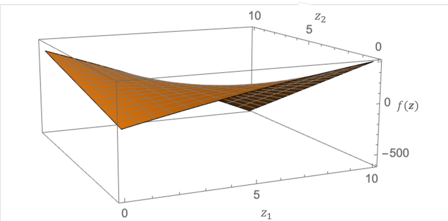


Mixed-integer nonlinear program.

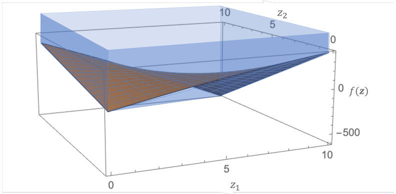
Characterize conv $\left(\mathcal{P}_f^{\mathcal{Z}(\mathcal{G},\mathbf{u})}\right)$

Theorem (full description of conv $\left(\mathcal{P}_f^{\mathcal{Z}(\mathcal{G},\mathbf{u})}\right)$) [Yu and Küçükyavuz, 2022]

The DR inequalities, MIR inequalities, along with the box and monotonicity constraints, fully describe conv $\left(\mathcal{P}_f^{\mathcal{Z}(\mathcal{G},\mathbf{u})}\right)$.



Mixed-integer nonlinear program.



Continuous linear program (with exponentially many constraints).

Exact Separation of DR Inequalities

- We propose an **exact separation** algorithm for DR inequalities.

Exact Separation of DR Inequalities

- We propose an **exact separation** algorithm for DR inequalities.
  - Finds a **most-violated** DR inequality at **any** $(\bar{\mathbf{z}}, \overline{w}) \notin \operatorname{conv}(\mathcal{Z}(\mathcal{G}, \mathbf{u})) \times \mathbb{R}$.

Exact Separation of DR Inequalities

- We propose an **exact separation** algorithm for DR inequalities.
  - Finds a **most-violated** DR inequality at **any** $(\bar{\mathbf{z}}, \overline{w}) \notin \operatorname{conv}(\mathcal{Z}(\mathcal{G}, \mathbf{u})) \times \mathbb{R}$.
  - An $\mathcal{O}(|\mathcal{V}|^2 \log |\mathcal{V}|)$ algorithm ($|\mathcal{V}|$ rounds of sorting).

Exact Separation of DR Inequalities

- We propose an **exact separation** algorithm for DR inequalities.
  - Finds a **most-violated** DR inequality at **any** $(\bar{\mathbf{z}}, \overline{w}) \notin \text{conv}\,(\mathcal{Z}(\mathcal{G}, \mathbf{u})) \times \mathbb{R}$.
  - An $\mathcal{O}(|\mathcal{V}|^2 \log |\mathcal{V}|)$ algorithm ($|\mathcal{V}|$ rounds of sorting).
- $\min_{\mathbf{z} \in \mathcal{Z}(\mathcal{G}, \mathbf{u})} f(\mathbf{z})$ is **polynomial-time solvable!**

Takeaways

- DR-submodular optimization is a **mixed-integer extension** of classical submodular optimization.
- A polyhedral study on DR-submodular minimization
  - under **box** and possibly additional **monotonicity** constraints,
  - with **mixed-integer** variables.
- Propose **valid linear inequalities** and the complete **convex hull description** for the epigraph.

  Nonlinear program $\rightarrow$ Linear program.
- Provide an **exact separation** algorithm.
- Establish **polynomial** time complexity of this class of constrained mixed-integer DR-submodular minimization problems.

Takeaways

- DR-submodular optimization is a **mixed-integer extension** of classical submodular optimization.
- A polyhedral study on DR-submodular minimization
  - under **box** and possibly additional **monotonicity** constraints,
  - with **mixed-integer** variables.
- Propose **valid linear inequalities** and the complete **convex hull description** for the epigraph.

  Nonlinear program $\rightarrow$ Linear program.
- Provide an **exact separation** algorithm.
- Establish **polynomial** time complexity of this class of constrained mixed-integer DR-submodular minimization problems.

Benign non-convexity!

References I

A. Atamtürk and V. Narayanan. Polymatroids and mean-risk minimization in discrete optimization. *Operations Research Letters*, 36(5):618–622, 2008.

A. Atamtürk and V. Narayanan. Submodular function minimization and polarity. *Mathematical Programming*, pages 1–11, 2021.

F. Bach. Submodular functions: from discrete to continuous domains. *Mathematical Programming*, 175(1):419–459, 2019.

A. A. Bian, B. Mirzasoleiman, J. Buhmann, and A. Krause. Guaranteed non-convex optimization: Submodular maximization over continuous domains. In *Artificial Intelligence and Statistics*, pages 111–120. PMLR, 2017.

J. Edmonds. Submodular functions, matroids, and certain polyhedra. In R. Guy, H. Hanani, N. Sauer, and J. Schönheim, editors, *Combinatorial Structures and Their Applications*, pages 69–87, New York, 1970. Gordon and Breach.

A. Ene and H. Nguyen. Parallel algorithm for non-monotone DR-submodular maximization. In *International Conference on Machine Learning*, pages 2902–2911. PMLR, 2020.

A. Ene and H. L. Nguyen. A reduction for optimizing lattice submodular functions with diminishing returns. *arXiv preprint arXiv:1606.08362*, 2016.

References II

S. Iwata, L. Fleischer, and S. Fujishige. A combinatorial strongly polynomial algorithm for minimizing submodular functions. *Journal of the ACM (JACM)*, 48(4):761–777, 2001.

L. Lovász. Submodular functions and convexity. In *Mathematical programming the state of the art*, pages 235–257. Springer, 1983.

H. Medal and I. Ahanor. A spatial branch-and-bound approach for maximization of non-factorable monotone continuous submodular functions. *arXiv preprint arXiv:2208.13886*, 2022.

G. L. Nemhauser and L. A. Wolsey. A recursive procedure to generate all cuts for 0–1 mixed integer programs. *Mathematical Programming*, 46(1):379–390, 1990.

R. Niazadeh, T. Roughgarden, and J. Wang. Optimal algorithms for continuous non-monotone submodular and DR-submodular maximization. *Advances in Neural Information Processing Systems*, 31, 2018.

J. B. Orlin. A faster strongly polynomial time algorithm for submodular function minimization. *Mathematical Programming*, 118(2):237–251, 2009.

References III

O. Sadeghi and M. Fazel. Online continuous DR-submodular maximization with long-term budget constraints. In *International Conference on Artificial Intelligence and Statistics*, pages 4410–4419. PMLR, 2020.

T. Soma and Y. Yoshida. Non-monotone DR-submodular function maximization. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 31, 2017.

T. Soma and Y. Yoshida. Maximizing monotone submodular functions over the integer lattice. *Mathematical Programming*, 172(1):539–563, 2018.

Z. Svitkina and L. Fleischer. Submodular approximation: Sampling-based algorithms and lower bounds. *SIAM Journal on Computing*, 40(6):1715–1737, 2011.

Q. Yu and S. Küçükyavuz. On constrained mixed-integer DR-submodular minimization. *arXiv preprint arXiv:2211.07726*, 2022.