Compact Formulations for Low-rank Functions with Indicator Variables

Shaoning Han

Department of Industrial & Systems Engineering
University of Southern California

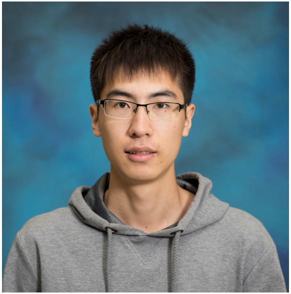
shaoning@usc.edu

INFORMS Annual Meeting, October 2023

Agenda

1. Introduction
2. Main results - convex hull description
3. Conclusions

Authors



Shaoning Han
Postdoctoral Researcher
ISE, USC

Andres Gomez
Assistant Professor
ISE, USC

Introduction

Consider

$$
\begin{aligned}
\min_{x,z} \ & \sum_k f_k(x) + a^\top x + c^\top z \\
\text{s.t. } & x_i(1 - z_i) = 0, \ z_i \in \{0, 1\} && \forall i \in [n] \\
& x_i \geq 0 && \forall i \in \mathcal{I}_+ \subseteq [n] \\
& \text{other constraints on } (x, z),
\end{aligned}
$$

where each f_k is convex.

Introduction

Consider

$$\begin{aligned} \min_{x,z} \ & \sum_k f_k(x) + a^\top x + c^\top z \\ \text{s.t.} \ & x_i(1 - z_i) = 0, \ z_i \in \{0, 1\} && \forall i \in [n] \\ & x_i \geq 0 && \forall i \in \mathcal{I}_+ \subseteq [n] \\ & \text{other constraints on } (x, z), \end{aligned}$$

where each f_k is convex.

- On/off constraints: $z_i = 0 \Rightarrow x_i = 0$

Introduction

Consider

$$
\begin{aligned}
\min_{x,z}\ & \sum_k f_k(x) + a^\top x + c^\top z \\
\text{s.t. }\ & x_i(1 - z_i) = 0,\ z_i \in \{0,1\} && \forall i \in [n] \\
& x_i \geq 0 && \forall i \in \mathcal{I}_+ \subseteq [n] \\
& \text{other constraints on } (x, z),
\end{aligned}
$$

where each f_k is convex.

- On/off constraints: $z_i = 0 \Rightarrow x_i = 0$

Introduction

Consider

$$\begin{aligned} \min_{x,z} \ & \sum_k f_k(x) + a^\top x + c^\top z \\ \text{s.t.} \ & x_i(1 - z_i) = 0, \ z_i \in \{0, 1\} && \forall i \in [n] \\ & x_i \geq 0 && \forall i \in \mathcal{I}_+ \subseteq [n] \\ & \text{other constraints on } (x, z), \end{aligned}$$

where each f_k is convex.

- On/off constraints: $z_i = 0 \Rightarrow x_i = 0$
- Nonconvex and $\mathcal{NP}$-hard

Introduction

Consider

$$\begin{aligned} \min_{x,z}\ & \sum_k f_k(x) + a^\top x + c^\top z \\ \text{s.t. } & x_i(1 - z_i) = 0,\ z_i \in \{0,1\} && \forall i \in [n] \\ & x_i \geq 0 && \forall i \in \mathcal{I}_+ \subseteq [n] \\ & \text{other constraints on } (x,z), \end{aligned}$$

where each f_k is convex.

- On/off constraints: $z_i = 0 \Rightarrow x_i = 0$
- Nonconvex and $\mathcal{NP}$-hard

Assumption Each f_k is a low-rank closed convex function

Introduction

Assumption Each f_k is a low-rank closed convex function

Definition (Rank; Rockafellar (1970))

The rank of f is the smallest integer k such that $f(x) = g(Ax) + c^\top x$ for some convex function $g : \mathbb{R}^k \to \mathbb{R}$ and matrix $A \in \mathbb{R}^{k \times n}$

Introduction

Assumption Each f_k is a low-rank closed convex function

Definition (Rank; Rockafellar (1970))

The rank of f is the smallest integer k such that $f(x) = g(Ax) + c^\top x$ for some convex function $g : \mathbb{R}^k \to \mathbb{R}$ and matrix $A \in \mathbb{R}^{k \times n}$

Examples

- $f(x) = c^\top x$, then $\text{rank}(f) = 0$

Introduction

Assumption Each f_k is a low-rank closed convex function

Definition (Rank; Rockafellar (1970))

The rank of f is the smallest integer k such that $f(x) = g(Ax) + c^\top x$ for some convex function $g : \mathbb{R}^k \to \mathbb{R}$ and matrix $A \in \mathbb{R}^{k \times n}$

Examples

- $f(x) = c^\top x$, then $\mathrm{rank}(f) = 0$
- $f(x) = \log\left(\exp(a^\top x) + 1\right) + c^\top x$, then $\mathrm{rank}(f) = 1$

Introduction

Assumption Each f_k is a low-rank closed convex function

Definition (Rank; Rockafellar (1970))

The rank of f is the smallest integer k such that $f(x) = g(Ax) + c^\top x$ for some convex function $g : \mathbb{R}^k \to \mathbb{R}$ and matrix $A \in \mathbb{R}^{k \times n}$

Examples

- $f(x) = c^\top x$, then $\operatorname{rank}(f) = 0$
- $f(x) = \log\left(\exp(a^\top x) + 1\right) + c^\top x$, then $\operatorname{rank}(f) = 1$
- $f(x) = g\left(a^\top x\right) + g\left(b^\top x\right)$, then $\operatorname{rank}(f) = 2$

Introduction

Assumption Each f_k is a low-rank closed convex function

Definition (Rank; Rockafellar (1970))

The rank of f is the smallest integer k such that $f(x) = g(Ax) + c^\top x$ for some convex function $g : \mathbb{R}^k \to \mathbb{R}$ and matrix $A \in \mathbb{R}^{k \times n}$

Examples

- $f(x) = c^\top x$, then $\operatorname{rank}(f) = 0$
- $f(x) = \log\left(\exp(a^\top x) + 1\right) + c^\top x$, then $\operatorname{rank}(f) = 1$
- $f(x) = g\left(a^\top x\right) + g\left(b^\top x\right)$, then $\operatorname{rank}(f) = 2$
- $f(x) = x^\top Q x + c^\top x$, then $\operatorname{rank}(f) = \operatorname{rank}(Q)$

Motivating application I – portfolio optimization

Portfolio index tracking problem Construct a portfolio of securities to reproduce the performance of a stock market index

$$\min_{x,z} \ (x - x_B)^\top Q (x - x_B)$$
$$\text{s.t. } x \geq 0, \ \sum_{i \in [n]} x_i = 1$$
$$\|x\|_0 \leq m$$

- $x_B \in \mathbb{R}^n$: benchmark index portfolio
- Q: covariance matrix of security returns
- m: maximum number of securities in the portfolio

Motivating application I – portfolio optimization

Portfolio index tracking problem Construct a portfolio of securities to reproduce the performance of a stock market index

$$\begin{aligned} \min_{x,z} \ & (x - x_B)^\top Q(x - x_B) \\ \text{s.t. } & x \geq 0, \ \sum_{i \in [n]} x_i = 1 \\ & \|x\|_0 \leq m \end{aligned}$$

- $x_B \in \mathbb{R}^n$: benchmark index portfolio
- Q: covariance matrix of security returns
- m: maximum number of securities in the portfolio

Motivating application I – portfolio optimization

Portfolio index tracking problem Construct a portfolio of securities to reproduce the performance of a stock market index

$$\begin{aligned} \min_{x,z}\ & (x - x_B)^\top Q(x - x_B) \\ \text{s.t. } & x \geq 0,\ \sum_{i \in [n]} x_i = 1 \\ & \|x\|_0 \leq m \end{aligned}$$

- $x_B \in \mathbb{R}^n$: benchmark index portfolio
- Q: covariance matrix of security returns
- m: maximum number of securities in the portfolio

Motivating application I – portfolio optimization

MIQP reformulation

$$
\begin{aligned}
\min_{x,z}\ & (x - x_B)^\top Q(x - x_B) \\
\text{s.t.}\ & x \geq 0,\ \sum_{i\in[n]} x_i = 1 \\
& x_i(1 - z_i) = 0,\ z_i \in \{0, 1\}\ \forall i \in [n] \\
& \sum_{i\in[n]} z_i \leq m
\end{aligned}
$$

- $z_i = 0 \Rightarrow x_i = 0$

Motivating application I – portfolio optimization

MIQP reformulation

$$\begin{aligned}\min_{x,z}\ & (x - x_B)^\top Q(x - x_B)\\ \text{s.t.}\ & x \geq 0,\ \sum_{i\in[n]} x_i = 1\\ & x_i(1 - z_i) = 0,\ z_i \in \{0, 1\}\ \forall i \in [n]\\ & \sum_{i\in[n]} z_i \leq m\end{aligned}$$

- $z_i = 0 \Rightarrow x_i = 0$

Motivating application I – portfolio optimization

MIQP reformulation

$$
\begin{aligned}
\min_{x,z}\ & (x - x_B)^\top Q(x - x_B) \\
\text{s.t. } & x \geq 0,\ \sum_{i \in [n]} x_i = 1 \\
& x_i(1 - z_i) = 0,\ z_i \in \{0, 1\}\ \forall i \in [n] \\
& \sum_{i \in [n]} z_i \leq m
\end{aligned}
$$

- $z_i = 0 \Rightarrow x_i = 0$

Motivating application I – portfolio optimization

MIQP reformulation

$$\begin{aligned} \min_{x,z}\ & (x - x_B)^\top Q (x - x_B) \\ \text{s.t. } & x \geq 0,\ \sum_{i \in [n]} x_i = 1 \\ & x_i(1 - z_i) = 0,\ z_i \in \{0,1\}\ \forall i \in [n] \\ & \sum_{i \in [n]} z_i \leq m \end{aligned}$$

- $z_i = 0 \Rightarrow x_i = 0$
- Covariances are estimated from a factor model (Bienstock (1996))

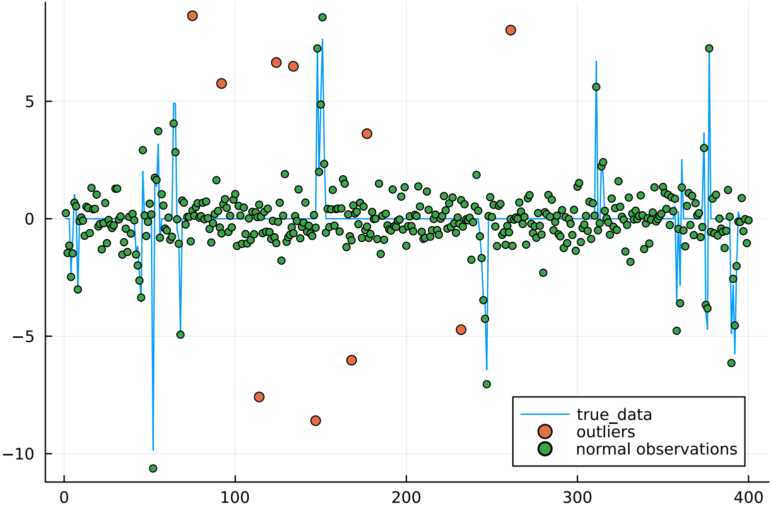
$$Q = FF^\top,$$

where $F \in \mathbb{R}^n \times \mathbb{R}^k,\ k \leq 20$ is small

Motivation application II – signal denoising

Signal denoising problem Given the noisy observations $c \in \mathbb{R}^n$ of a temporal process, consider

- Smooth
- Sparse
- Outliers



Motivation application II – signal denoising

Signal denoising problem Given the noisy observations $c \in \mathbb{R}^n$ of a temporal process, consider

$$\min_{x,v,z,w} \sum_{i=1}^{n} \underbrace{(x_i \qquad - c_i)^2}_{\text{fitness}} + \Omega \sum_{i=\ell+1}^{n} \underbrace{\left(x_i - \sum_{j=1}^{\ell} \alpha^j x_{i-\ell+j-1} \right)^2}_{\textit{smoothness}}$$

s.t.

- x_i: true values of the signal
- c_i: noisy observations

Motivation application II – signal denoising

Signal denoising problem Given the noisy observations $c \in \mathbb{R}^n$ of a temporal process, consider

$$\min_{x,v,z,w} \sum_{i=1}^{n} \underbrace{(x_i \qquad - c_i)^2}_{\text{fitness}} + \Omega \sum_{i=\ell+1}^{n} \underbrace{\left(x_i - \sum_{j=1}^{\ell} \alpha^j x_{i-\ell+j-1} \right)^2}_{\textit{smoothness}}$$

s.t.

- $\alpha = 0.9$: decaying factor of proximity
- Ω: weight of smoothness

Motivation application II – signal denoising

Signal denoising problem Given the noisy observations $c \in \mathbb{R}^n$ of a temporal process, consider

$$\min_{x,v,z,w} \sum_{i=1}^{n} \underbrace{(x_i \qquad - c_i)^2}_{\text{fitness}} + \Omega \sum_{i=\ell+1}^{n} \underbrace{\left(x_i - \sum_{j=1}^{\ell} \alpha^j x_{i-\ell+j-1} \right)^2}_{smoothness}$$

$$\text{s.t. } x_i(1 - z_i) = 0, \; z_i \in \{0,1\} \; \forall i \in [n], \; \sum_{i=1}^{n} z_i \leq k_1 \qquad \text{(sparsity)}$$

- $\alpha = 0.9$: decaying factor of proximity
- Ω: weight of smoothness

Motivation application II – signal denoising

Signal denoising problem Given the noisy observations $c \in \mathbb{R}^n$ of a temporal process, consider

$$\min_{x,v,z,w} \sum_{i=1}^{n} \underbrace{(x_i - v_i - c_i)^2}_{\text{fitness+robustness}} + \Omega \sum_{i=\ell+1}^{n} \underbrace{\left(x_i - \sum_{j=1}^{\ell} \alpha^j x_{i-\ell+j-1} \right)^2}_{smoothness}$$

$$\text{s.t. } x_i(1 - z_i) = 0,\ z_i \in \{0,1\}\ \forall i \in [n],\ \sum_{i=1}^{n} z_i \le k_1 \qquad \text{(sparsity)}$$

$$v_i(1 - w_i) = 0,\ w_i \in \{0,1\}\ \forall i \in [n],\ \sum_{i=1}^{n} w_i \le k_2 \qquad \text{(outlier)}$$

- $w_i = 0$: $v_i = 0 \Rightarrow c_i$ is not an outlier
- $w_i = 1$: $x_i - v_i - c_i = 0 \Rightarrow c_i$ is an outlier

Introduction

MINLP with low-rank structure

$$\begin{aligned}
\min_{x,z} \ & \sum_k f_k(x) + a^\top x + c^\top z \\
\text{s.t. } & x_i(1 - z_i) = 0,\ z_i \in \{0, 1\} && \forall i \in [n] \\
& x_i \geq 0 && \forall i \in \mathcal{I}_+ \subseteq [n] \\
& \text{other constraints on } (x, z),
\end{aligned}$$

To solve it efficiently, we study

$$\mathcal{Q} = \left\{ (t, x, z) \in \mathbb{R}^{n+1} \times \{0, 1\}^n : \begin{array}{r} t \geq f(x), x_i \geq 0 \ \forall i \in \mathcal{I}_+ \\ x_i(1 - z_i) = 0 \ \forall i \in [n] \end{array} \right\}$$

Goal: Compute $\operatorname{cl}\operatorname{conv} \mathcal{Q}$

Literature review

Known cases for $\text{cl conv}(\mathcal{Q})$

$$\mathcal{Q} = \left\{ (t, x, z) \in \mathbb{R}^{n+1} \times \{0,1\}^n : \begin{array}{c} t \geq f(x), x_i \geq 0 \ \forall i \in \mathcal{I}_+ \\ x_i(1 - z_i) = 0 \ \forall i \in [n] \end{array} \right\}$$

- $k = \text{rank}(f)$, Quad $=$ Quadratic, Conv $=$ General Convex

k or n	f	$\mathcal{I}_+ = \emptyset$ (x free)	$\mathcal{I}_+ = [n]$ ($x \geq 0$)
$n = 1$		Ceria et al.(1999), Frangioni et al.(2006), Aktürk et al.(2009), etc.	
$k = 1$	Quad	Atamtürk et al. (2019)	Atamtürk et al. (2023)
	Conv	Wei et al. (2022)	Shafieezadeh-Abadeh et al. (2023)
$n = 2$	Quad	?	Han et al. (2023), De Rosa et al. (2023)
$k \geq 2$		?	?

Literature review

Known cases for $\operatorname{cl}\operatorname{conv}(\mathcal{Q})$

$$\mathcal{Q} = \left\{ (t, x, z) \in \mathbb{R}^{n+1} \times \{0,1\}^n : \begin{array}{c} t \geq f(x), x_i \geq 0 \ \forall i \in \mathcal{I}_+ \\ x_i(1 - z_i) = 0 \ \forall i \in [n] \end{array} \right\}$$

- $k = \operatorname{rank}(f)$, Quad = Quadratic, Conv = General Convex

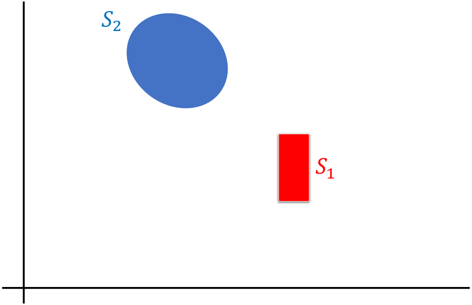
k or n	f	$\mathcal{I}_+ = \emptyset$ (x free)	$\mathcal{I}_+ = [n]$ ($x \geq 0$)
$n = 1$		Ceria et al.(1999), Frangioni et al.(2006), Aktürk et al.(2009), etc.	
$k = 1$	Quad	Atamtürk et al. (2019)	Atamtürk et al. (2023)
	Conv	Wei et al. (2022)	Shafieezadeh-Abadeh et al. (2023)
$n = 2$	Quad	✓	Han et al. (2023), De Rosa et al. (2023)
$k \geq 2$		✓	✓

This work ✓: $\underbrace{\text{Compact}}_{\mathcal{O}(n^k)}$ description of $\operatorname{cl}\operatorname{conv}(\mathcal{Q})$ using disjunctive programming

Preliminaries: disjunctive programming

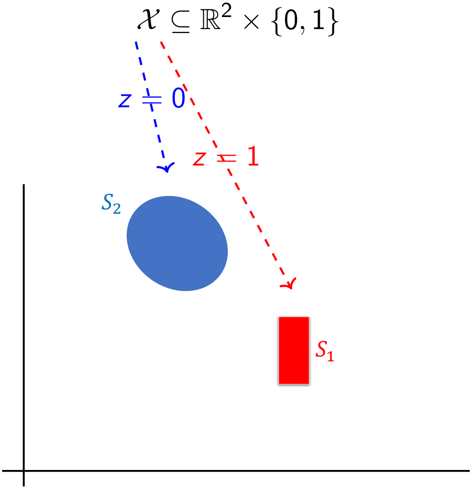
Assume one binary variable

$$\mathcal{X} \subseteq \mathbb{R}^2 \times \{0, 1\}$$



Preliminaries: disjunctive programming

Assume one binary variable

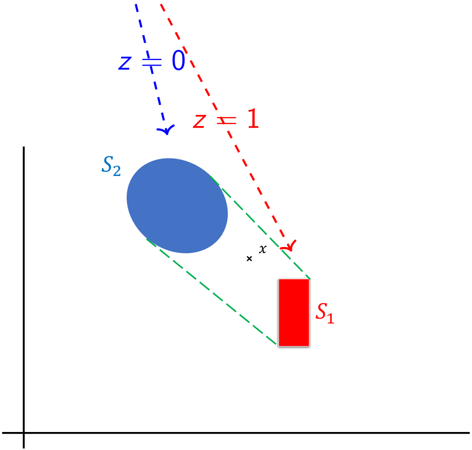


Preliminaries: disjunctive programming

Assume one binary variable

$$\mathcal{X} \subseteq \mathbb{R}^2 \times \{0, 1\}$$

Convex hull of $\mathcal{X}$

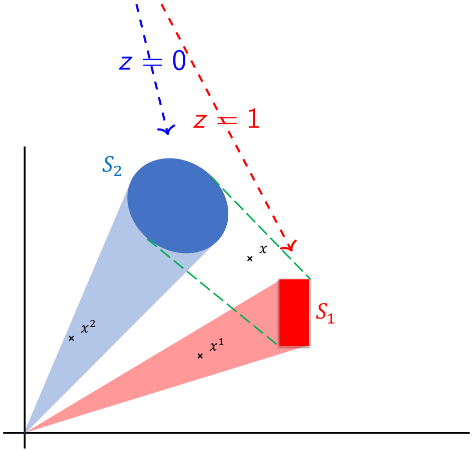


Preliminaries: disjunctive programming

Assume one binary variable

$$\mathcal{X} \subseteq \mathbb{R}^2 \times \{0,1\}$$

Convex hull of $\mathcal{X}$

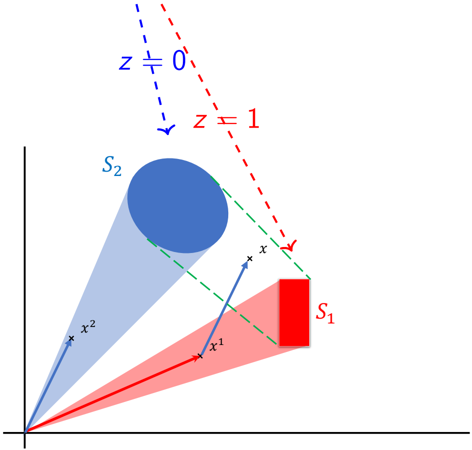


Preliminaries: disjunctive programming

Assume one binary variable

$$\mathcal{X} \subseteq \mathbb{R}^2 \times \{0, 1\}$$

Convex hull of $\mathcal{X}$



Preliminaries: disjunctive programming

For any mixed-binary set $\mathcal{X} \subseteq \mathbb{R}^m \times \{0,1\}^n$,

$$\mathcal{X} = \bigcup_{\bar{z} \in \{0,1\}^n} [\mathcal{X} \cap (\mathbb{R}^m \times \{\bar{z}\})]$$

- # of disjunctions $= 2^n$
- Disjunctive programming $\Rightarrow$ describe conv $\mathcal{X}$ in a lifted space

$$\# \text{ of additional vars} \approx \dim(\mathcal{X}) \times \# \text{ of disjunctions} = \mathcal{O}((n+m)2^n)$$

$\Rightarrow$ Only applicable in practice for small n

Preliminaries: disjunctive programming

For any mixed-binary set $\mathcal{X} \subseteq \mathbb{R}^m \times \{0,1\}^n$,

$$\mathcal{X} = \bigcup_{\bar{z} \in \{0,1\}^n} [\mathcal{X} \cap (\mathbb{R}^m \times \{\bar{z}\})]$$

- $\#$ of disjunctions $= 2^n$
- Disjunctive programming $\Rightarrow$ describe $\operatorname{conv} \mathcal{X}$ in a lifted space

$$\# \text{ of additional vars} \approx \dim(\mathcal{X}) \times \# \text{ of disjunctions} = \mathcal{O}((n+m)2^n)$$

$\Rightarrow$ Only applicable in practice for small n

N/A to our setting:

$$\mathcal{Q} = \left\{ (t, x, z) \in \mathbb{R}^{n+1} \times \{0,1\}^n : \begin{array}{c} t \geq f(x), x_i \geq 0 \; \forall i \in \mathcal{I}_+ \\ x_i(1 - z_i) = 0 \; \forall i \in [n] \end{array} \right\}$$

Agenda

Homogeneous cases

Consider

$$\mathcal{Q} = \left\{ (t, x, z) \in \mathbb{R}^{n+1} \times \{0,1\}^n : \begin{array}{r} t \geq f(x), x_i \geq 0 \ \forall i \in \mathcal{I}_+ \\ x_i(1 - z_i) = 0 \ \forall i \in [n] \end{array} \right\},$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is convex and $\mathcal{I}_+ \subseteq [n]$

Homogeneous cases

Consider

$$\mathcal{Q} = \left\{ (t, x, z) \in \mathbb{R}^{n+1} \times \{0, 1\}^n : \begin{array}{c} t \geq f(x), x_i \geq 0 \ \forall i \in \mathcal{I}_+ \\ x_i(1 - z_i) = 0 \ \forall i \in [n] \end{array} \right\},$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is convex and $\mathcal{I}_+ \subseteq [n]$

Proposition (Han and Gómez (2021))

If f is positively homogeneous, i.e. $f(\lambda x) = \lambda f(x)$ for all $\lambda \geq 0$, then

$$\text{cl conv}\, \mathcal{Q} = \left\{ (t, x, z) \in \mathbb{R}^{n+1} \times [0, 1]^n : \begin{array}{c} t \geq f(x), x_i \geq 0 \ \forall i \in \mathcal{I}_+ \\ \cancel{x_i(1 - z_i) = 0 \ \forall i \in [n]} \end{array} \right\}.$$

New disjunctive representation in non-homogeneous cases

General (non-homogeneous) cases

Observe x_i and z_i are linked only through $x_i(1 - z_i) = 0$

New disjunctive representation in non-homogeneous cases

General (non-homogeneous) cases

Observe x_i and z_i are linked only through $x_i(1 - z_i) = 0$

$\Rightarrow$ decompose $\mathcal{Q}$ based on either $x_i = 0$ or $z_i = 1$

$$\mathcal{Q} = \bigcup_{\mathcal{I} \subseteq [n]} \mathcal{Q} \cap \{(t, x, z) : z_i = 1 \; \forall i \in \mathcal{I}, \; x_i = 0 \; \forall i \notin \mathcal{I}\}$$

New disjunctive representation in non-homogeneous cases

General (non-homogeneous) cases

Observe x_i and z_i are linked only through $x_i(1-z_i)=0$

$\Rightarrow$ decompose $\mathcal{Q}$ based on either $x_i=0$ or $z_i=1$

$$\begin{aligned}\mathcal{Q} &= \bigcup_{\mathcal{I}\subseteq[n]} \mathcal{Q} \cap \{(t,x,z) : z_i = 1\ \forall i \in \mathcal{I},\ x_i = 0\ \forall i \notin \mathcal{I}\} \\ &= \bigcup_{\mathcal{I}\subseteq[n]} \underbrace{\{(t,x) : t \geq f(x), x_i = 0\, \forall i \notin \mathcal{I}\}}_{\mathcal{X}(\mathcal{I})} \times \underbrace{\{z \in \{0,1\}^n : z_i = 1\, \forall i \in \mathcal{I}\}}_{\mathcal{Z}(\mathcal{I})}\end{aligned}$$

New disjunctive representation in non-homogeneous cases

General (non-homogeneous) cases

Observe x_i and z_i are linked only through $x_i(1 - z_i) = 0$

$\Rightarrow$ decompose $\mathcal{Q}$ based on either $x_i = 0$ or $z_i = 1$

$$\begin{aligned}
\mathcal{Q} &= \bigcup_{\mathcal{I} \subseteq [n]} \mathcal{Q} \cap \{(t, x, z) : z_i = 1 \, \forall i \in \mathcal{I}, \; x_i = 0 \, \forall i \notin \mathcal{I}\} \\
&= \bigcup_{\mathcal{I} \subseteq [n]} \underbrace{\{(t, x) : t \geq f(x), x_i = 0 \, \forall i \notin \mathcal{I}\}}_{\mathcal{X}(\mathcal{I})} \times \underbrace{\{z \in \{0, 1\}^n : z_i = 1 \, \forall i \in \mathcal{I}\}}_{\mathcal{Z}(\mathcal{I})} \\
&=: \bigcup_{\mathcal{I} \subseteq [n]} \underbrace{\mathcal{X}(\mathcal{I}) \times \mathcal{Z}(\mathcal{I})}_{\mathcal{V}(\mathcal{I})}
\end{aligned}$$

New disjunctive representation in non-homogeneous cases

General (non-homogeneous) cases

Observe x_i and z_i are linked only through $x_i(1 - z_i) = 0$

$\Rightarrow$ decompose $\mathcal{Q}$ based on either $x_i = 0$ or $z_i = 1$

$$\begin{aligned}
\mathcal{Q} &= \bigcup_{\mathcal{I} \subseteq [n]} \mathcal{Q} \cap \{(t, x, z) : z_i = 1 \; \forall i \in \mathcal{I}, \; x_i = 0 \; \forall i \notin \mathcal{I}\} \\
&= \bigcup_{\mathcal{I} \subseteq [n]} \underbrace{\{(t, x) : t \geq f(x), x_i = 0 \; \forall i \notin \mathcal{I}\}}_{\mathcal{X}(\mathcal{I})} \times \underbrace{\{z \in \{0,1\}^n : z_i = 1 \; \forall i \in \mathcal{I}\}}_{\mathcal{Z}(\mathcal{I})} \\
&=: \bigcup_{\mathcal{I} \subseteq [n]} \underbrace{\mathcal{X}(\mathcal{I}) \times \mathcal{Z}(\mathcal{I})}_{\mathcal{V}(\mathcal{I})}
\end{aligned}$$

Note $\mathcal{X}(\mathcal{I})$ is convex and $\operatorname{conv} \mathcal{Z}(\mathcal{I}) = \{z \in [0,1]^n : z_i = 1 \; \forall i \in \mathcal{I}\}$

New disjunctive representation in non-homogeneous cases

General (non-homogeneous) cases

Observe x_i and z_i are linked only through $x_i(1-z_i)=0$

$\Rightarrow$ decompose $\mathcal{Q}$ based on either $x_i=0$ or $z_i=1$

$$\begin{aligned}\mathcal{Q} &= \bigcup_{\mathcal{I}\subseteq[n]} \mathcal{Q}\cap\{(t,x,z) : z_i = 1\ \forall i\in\mathcal{I},\ x_i = 0\ \forall i\notin\mathcal{I}\} \\ &= \bigcup_{\mathcal{I}\subseteq[n]} \underbrace{\{(t,x) : t\geq f(x), x_i = 0\ \forall i\notin\mathcal{I}\}}_{\mathcal{X}(\mathcal{I})} \times \underbrace{\{z\in\{0,1\}^n : z_i = 1\ \forall i\in\mathcal{I}\}}_{\mathcal{Z}(\mathcal{I})} \\ &=: \bigcup_{\mathcal{I}\subseteq[n]} \underbrace{\mathcal{X}(\mathcal{I})\times\mathcal{Z}(\mathcal{I})}_{\mathcal{V}(\mathcal{I})}\end{aligned}$$

Note $\mathcal{X}(\mathcal{I})$ is convex and $\operatorname{conv}\mathcal{Z}(\mathcal{I}) = \{z\in[0,1]^n : z_i = 1\ \forall i\in\mathcal{I}\}$

Still 2^n disjunctions! $\Rightarrow$

New disjunctive representation in non-homogeneous cases

General (non-homogeneous) cases

Observe x_i and z_i are linked only through $x_i(1 - z_i) = 0$

$\Rightarrow$ decompose $\mathcal{Q}$ based on either $x_i = 0$ or $z_i = 1$

$$\begin{aligned}
\mathcal{Q} &= \bigcup_{\mathcal{I} \subseteq [n]} \mathcal{Q} \cap \{(t, x, z) : z_i = 1 \ \forall i \in \mathcal{I}, \ x_i = 0 \ \forall i \notin \mathcal{I}\} \\
&= \bigcup_{\mathcal{I} \subseteq [n]} \underbrace{\{(t, x) : t \geq f(x), x_i = 0 \, \forall i \notin \mathcal{I}\}}_{\mathcal{X}(\mathcal{I})} \times \underbrace{\{z \in \{0, 1\}^n : z_i = 1 \, \forall i \in \mathcal{I}\}}_{\mathcal{Z}(\mathcal{I})} \\
&=: \bigcup_{\mathcal{I} \subseteq [n]} \underbrace{\mathcal{X}(\mathcal{I}) \times \mathcal{Z}(\mathcal{I})}_{\mathcal{V}(\mathcal{I})}
\end{aligned}$$

Note $\mathcal{X}(\mathcal{I})$ is convex and $\operatorname{conv} \mathcal{Z}(\mathcal{I}) = \{z \in [0, 1]^n : z_i = 1 \ \forall i \in \mathcal{I}\}$

Still 2^n disjunctions! $\Rightarrow$ Exploit the low-rank structure $f(x) = g(Ax)$

Convex hull description of $\mathcal{Q}$

Theorem (Han and Gómez (2021))

Assume $\operatorname{rank}(f) \leq k$ *and* $f(0) = 0$. *Then*

$$\operatorname{cl}\operatorname{conv}(\mathcal{Q}) = \operatorname{cl}\operatorname{conv}\left(\left(\bigcup_{\mathcal{I}:|\mathcal{I}|\leq k} \mathcal{V}(\mathcal{I}) \cup \mathcal{R}\right)\right),$$

where

$$\mathcal{R} = \{(t, x, z) : t \geq 0, Ax = 0, x_i \geq 0,\ \forall i \in \mathcal{I}_+, z_i = 1,\ \forall i \in [n]\}$$

- $\mathcal{V}(\mathcal{I})$: "extreme points" of $\operatorname{cl}\operatorname{conv}\mathcal{Q}$
- $\mathcal{R}$: "extreme rays" of $\operatorname{cl}\operatorname{conv}\mathcal{Q}$
- $\mathcal{O}(n^k)$ number of disjunctions

Convex hull description of $\mathcal{Q}$

Proof outline.

Consider

$$\begin{aligned} \min\ & a^\top x + c^\top z + g(Ax) \\ \text{s.t. } & x_i \geq 0 \quad \forall i \in \mathcal{I}_+ \\ & x \circ (1 - z) = 0 \end{aligned} \tag{MINLP}$$

Assume $(\bar{x}, \bar{z})$ is the optimal solution to (MINLP).

Convex hull description of $\mathcal{Q}$

Proof outline.

Consider

$$\begin{aligned} \min\ & a^\top x + c^\top z + g(Ax) \\ \text{s.t. }\ & x_i \geq 0 \quad \forall i \in \mathcal{I}_+ \\ & x \circ (1 - z) = 0 \end{aligned} \tag{MINLP}$$

Assume $(\bar{x}, \bar{z})$ is the optimal solution to (MINLP). Then

$$\begin{aligned} \min\ & a^\top x + g(A\bar{x}) \\ \text{s.t. }\ & Ax = A\bar{x} \\ & \bar{x}_i x_i \geq 0 \quad \forall i \\ & x_i = 0 \quad \forall i : \bar{z}_i = 0 \end{aligned} \tag{LP}$$

has an optimal solution $\hat{x}$ with at most $\text{rank}(A) = k$ nonzero entries. Moreover, $(\hat{x}, \bar{z})$ is optimal to (MINLP).

Convex hull description of $\mathcal{Q}$

Proof outline.

Consider

$$\begin{aligned} \min\ & a^\top x + c^\top z + g(Ax) \\ \text{s.t.}\ & x_i \geq 0 \quad \forall i \in \mathcal{I}_+ \\ & x \circ (1 - z) = 0 \end{aligned} \tag{MINLP}$$

Assume $(\bar{x}, \bar{z})$ is the optimal solution to (MINLP). Then

$$\begin{aligned} \min\ & a^\top x + g(A\bar{x}) \\ \text{s.t.}\ & Ax = A\bar{x} \\ & \bar{x}_i x_i \geq 0 \quad \forall i \\ & x_i = 0 \quad \forall i : \bar{z}_i = 0 \end{aligned} \tag{LP}$$

has an optimal solution $\hat{x}$ with at most $\text{rank}(A) = k$ nonzero entries.
Moreover, $(\hat{x}, \bar{z})$ is optimal to (MINLP).
$\Rightarrow (\hat{x}, \bar{z}) \in$ a certain $\mathcal{V}(\mathcal{I})$ with $|\mathcal{I}| \leq k$. □

Implementation: perspective function

Definition (Perspective function)

Given a closed convex function $f : \mathbb{R}^n \to \mathbb{R}$, its perspective function $f^\pi(x, \lambda)$ is defined as

$$f^\pi(x,\lambda) = \begin{cases} \lambda f\left(\frac{x}{\lambda}\right) & \text{if } \lambda > 0 \\ \lim\limits_{\lambda \to 0} \lambda f\left(\frac{x}{\lambda}\right) & \text{if } \lambda = 0 \\ +\infty & \text{o.w.} \end{cases}$$

- f^π is closed, convex, positively homogeneous

Implementation: perspective function

Definition (Perspective function)

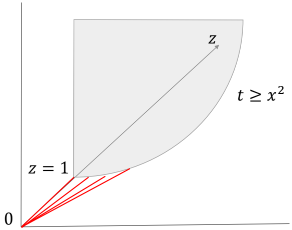
Given a closed convex function $f : \mathbb{R}^n \to \mathbb{R}$, its perspective function $f^\pi(x, \lambda)$ is defined as

$$f^\pi(x,\lambda) = \begin{cases} \lambda f\left(\frac{x}{\lambda}\right) & \text{if } \lambda > 0 \\ \lim\limits_{\lambda \to 0} \lambda f\left(\frac{x}{\lambda}\right) & \text{if } \lambda = 0 \\ +\infty & \text{o.w.} \end{cases}$$

- f^π is closed, convex, positively homogeneous

Example $f(x) = x^2$

$$\Rightarrow f^\pi(x,\lambda) = \begin{cases} x^2/\lambda & \text{if } \lambda > 0 \\ 0 & \text{if } (x,\lambda) = 0 \\ +\infty & \text{o.w.} \end{cases}$$



Case study

Rank-one case $f = g\left(\sum_{i=1}^{n} a_i x_i\right)$

- New DP representation + low rank ⇒

Proposition (Han and Gómez (2021))

Point $(t, x, z) \in \operatorname{cl}\operatorname{conv} \mathcal{Q}$ *if and only if there exists* $\lambda, \tau \in \mathbb{R}^n$ *such that the following inequality system is consistent*

$$
\begin{aligned}
& t \geq \sum_{i=1}^{n} g^{\pi}(a_i(x_i - \tau_i), \lambda_i), \\
& a^{\top}\tau = 0,\ 0 \leq \tau_i \leq x_i\ \forall i \in \mathcal{I}_+, \\
& \lambda_i \leq z_i \leq 1\ \forall i \in [n], \\
& \lambda \geq 0,\ \sum_{i=1}^{n} \lambda_i \leq 1
\end{aligned}
$$

More discussion in rank-one case

k or n	f	$\mathcal{I}_+ = \emptyset$ (x free)	$\mathcal{I}_+ = [n]$ ($x \geq 0$)
$n = 1$		Ceria et al.(1999), Frangioni et al.(2006), Aktürk et al.(2009), etc.	
$k = 1$	Quad	Atamtürk et al. (2019)	Atamtürk et al. (2023)
	Conv	Wei et al. (2022)	Shafieezadeh-Abadeh et al. (2023)
$n = 2$	Quad	✓	Han et al. (2023), De Rosa et al. (2023)
$k \geq 2$		✓	✓

More discussion in rank-one case

k or n	f	$\mathcal{I}_+ = \emptyset$ (x free)	$\mathcal{I}_+ = [n]$ ($x \geq 0$)
$n = 1$		Ceria et al.(1999), Frangioni et al.(2006), Aktürk et al.(2009), etc.	
$k = 1$	Quad	Atamtürk et al. (2019)	Atamtürk et al. (2023)
	Conv	Wei et al. (2022)	Shafieezadeh-Abadeh et al. (2023)
$n = 2$	Quad	✓	Han et al. (2023), De Rosa et al. (2023)
$k \geq 2$		✓	✓

- Atamtürk et al. (2023): cl conv $\mathcal{Q}$ described by cutting surfaces with $\mathcal{O}(n)$ additional vars per cut
- Our results: more compact extended formulation ($\mathcal{O}(n)$ in total)

More discussion in rank-one case

k or n	f	$\mathcal{I}_+ = \emptyset$ (x free)	$\mathcal{I}_+ = [n]$ ($x \geq 0$)
$n = 1$		Ceria et al.(1999), Frangioni et al.(2006), Aktürk et al.(2009), etc.	
$k = 1$	Quad	Atamtürk et al. (2019)	Atamtürk et al. (2023)
	Conv	Wei et al. (2022)	Shafieezadeh-Abadeh et al. (2023)
$n = 2$	Quad	✓	Han et al. (2023), De Rosa et al. (2023)
$k \geq 2$		✓	✓

- Atamtürk et al. (2023): cl conv $\mathcal{Q}$ described by cutting surfaces with $\mathcal{O}(n)$ additional vars per cut
- Our results: more compact extended formulation ($\mathcal{O}(n)$ in total)

$\Rightarrow$ More efficient implementation

Experimental results - portfolio optimization

Cutting surface implementation v.s. Extended formulation

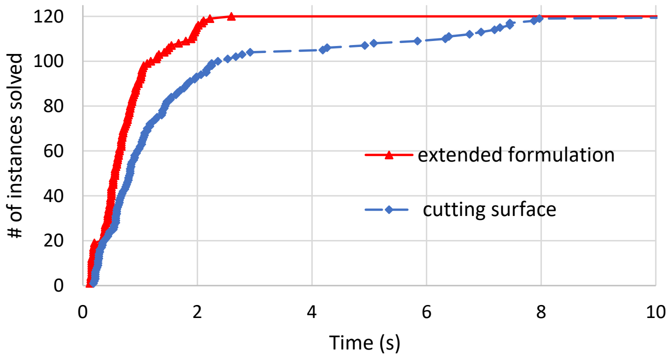


Figure: Number of instances solved as a function of time.

- Average time: cutting surface 1.79s v.s. extended formulation 0.78s
- Maximum time: cutting surface 13.3s v.s. extended formulation 2.6s

Experimental results - signal denoising

Rank-two v.s. Rank-one v.s. Big-M

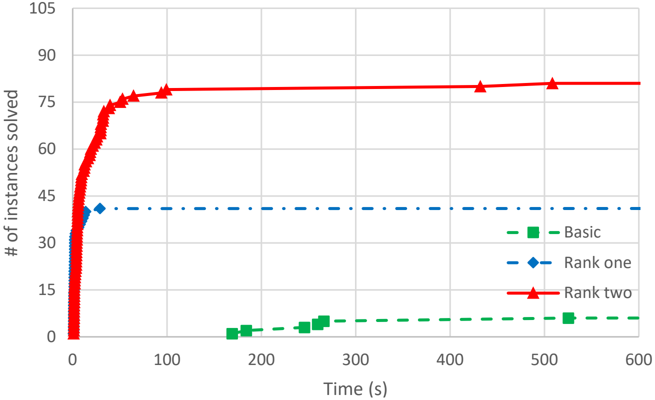


Figure: Number of instances solved as a function of time

Agenda

Take home message

- New DP representation for low-rank functions with indicator variables
- Compact extended formulation for convex hull description
- More efficient implementation in practice

Our paper is available at: `https://arxiv.org/abs/2110.14884`

Take home message

- New DP representation for low-rank functions with indicator variables
- Compact extended formulation for convex hull description
- More efficient implementation in practice

Our paper is available at: `https://arxiv.org/abs/2110.14884`

Thank You!

Reference I

Aktürk, M. S., Atamtürk, A., and Gürel, S. (2009). A strong conic quadratic reformulation for machine-job assignment with controllable processing times. Operations Research Letters, 37(3):187–191.

Atamtürk, A. and Gómez, A. (2019). Rank-one convexification for sparse regression. arXiv preprint arXiv:1901.10334.

Atamtürk, A. and Gómez, A. (2023). Supermodularity and valid inequalities for quadratic optimization with indicators. Mathematical Programming, 201(1-2):295–338.

Bienstock, D. (1996). Computational study of a family of mixed-integer quadratic programming problems. Mathematical Programming, 74(2):121–140.

Ceria, S. and Soares, J. (1999). Convex programming for disjunctive convex optimization. Mathematical Programming, 86:595–614.

De Rosa, A. and Khajavirad, A. (2022). Explicit convex hull description of bivariate quadratic sets with indicator variables. arXiv preprint arXiv:2208.08703.

Reference II

Frangioni, A. and Gentile, C. (2006). Perspective cuts for a class of convex 0–1 mixed integer programs. Mathematical Programming, 106(2):225–236.

Han, S. and Gómez, A. (2021). Compact extended formulations for low-rank functions with indicator variables. Submitted to Mathematics of Operations Research.

Han, S., Gómez, A., and Atamtürk, A. (2023). 2×2-convexifications for convex quadratic optimization with indicator variables. Mathematical Programming.

Rockafellar, R. T. (1970). Convex analysis, volume 18. Princeton university press.

Shafieezadeh-Abadeh, S. and Kılınç-Karzan, F. (2023). Constrained optimization of rank-one functions with indicator variables. arXiv preprint arXiv:2303.18158.

Wei, L., Gómez, A., and Küçükyavuz, S. (2022). Ideal formulations for constrained convex optimization problems with indicator variables. Mathematical Programming, 192(1):57–88.