Unifying Branch-and-Bound Approaches to Solve L0-Penalized Problems

Theo Guyard with C. Herzet, A. Arslan and C. Elvira

Inria Rennes, France

SIAM Conference on Optimization
Seattle, USA
2023

L0-penalized problems

Sparse optimization

Minimize a loss ⟷ **Sparse optimization** ⟷ Sparse solution

Machine Learning
Signal processing
Network design
...

Minimize a loss ⟷ **Sparse optimization** ⟷ **Sparse solution**

Machine Learning
Signal processing
Network design
...

ℓ_0-penalized problem

$$\min_x \; F(x) + \lambda \|x\|_0 + G(x)$$

$$\begin{aligned} F(\cdot) &: \text{loss} \\ \|\cdot\|_0 &: \text{sparsity} \\ \lambda &: \text{trade-off} \\ G(\cdot) &: \text{modelling} \end{aligned}$$

Contributions

L0-penalized problem

$\min_x \ F(x) + \lambda \|x\|_0 + G(x)$

NP-hard

Contributions

L0-penalized problem

$$\min_{x} \; F(x) + \lambda\|x\|_0 + G(x)$$

NP-hard

Generic approaches

- D. Bertsimas (2016)
- S. Bourguignon (2017)
- A. Atamtürk (2020)
- D. Bertsimas (2021)
- C. Kanzow (2022)
- ...

✓ Off-the-shelf MIP solver
✗ Slow
✓ Arbitrary $F(\cdot)$ and $G(\cdot)$

Contributions

L0-penalized problem

$$\min_x \; F(x) + \lambda\|x\|_0 + G(x)$$

NP-hard

Generic approaches

- D. Bertsimas (2016)
- S. Bourguignon (2017)
- A. Atamtürk (2020)
- D. Bertsimas (2021)
- C. Kanzow (2022)
- ...

✓ Off-the-shelf MIP solver
✗ Slow
✓ Arbitrary $F(\cdot)$ and $G(\cdot)$

Tailored approaches

- R. Ben Mhenni (2021)
- H. Hazimeh (2021)
- G. Samain (2022)
- A. Olama (2022)
- A. Atamtürk (2022)
- ...

✓ Branch-and-Bound
✓ Fast
✗ Specific $F(\cdot)$ and $G(\cdot)$

Contributions

L0-penalized problem

$$\min_x \; F(x) + \lambda\|x\|_0 + G(x)$$

NP-hard

Generic approaches

- D. Bertsimas (2016)
- S. Bourguignon (2017)
- A. Atamtürk (2020)
- D. Bertsimas (2021)
- C. Kanzow (2022)
- ...

✓ Off-the-shelf MIP solver
✗ Slow
✓ Arbitrary $F(\cdot)$ and $G(\cdot)$

Tailored approaches

- R. Ben Mhenni (2021)
- H. Hazimeh (2021)
- G. Samain (2022)
- A. Olama (2022)
- A. Atamtürk (2022)
- ...

✓ Branch-and-Bound
✓ Fast
✗ Specific $F(\cdot)$ and $G(\cdot)$

Contribution

- This talk (2023)
- Extended paper (202?)

✓ Branch-and-Bound
✓ Fast
✓ Arbitrary $F(\cdot)$ and $G(\cdot)$

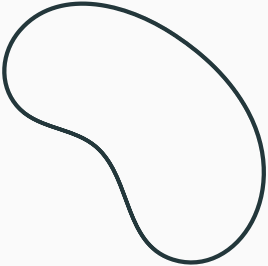
Branch-and-Bound

Algorithmic concept

"Enumerate all candidate solutions and discard sub-optimal ones."

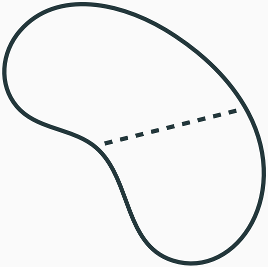
Algorithmic concept

"Enumerate all candidate solutions and discard sub-optimal ones."



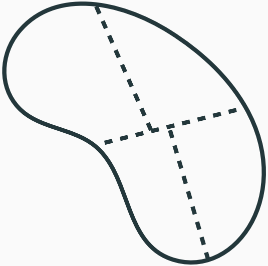
Algorithmic concept

"Enumerate all candidate solutions and discard sub-optimal ones."



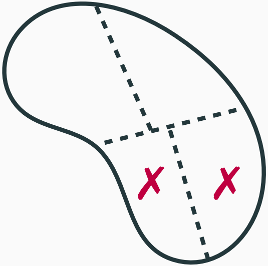
Algorithmic concept

"Enumerate all candidate solutions and discard sub-optimal ones."



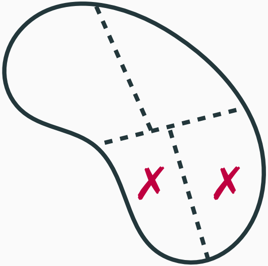
Algorithmic concept

"Enumerate all candidate solutions and discard sub-optimal ones."



Algorithmic concept

"Enumerate all candidate solutions and discard sub-optimal ones."



Main principles

Branching: Divide the search space

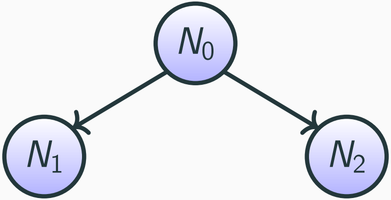
Bounding: Test whether a region can contain optimal solutions

Pruning: Discard regions without optimal solutions

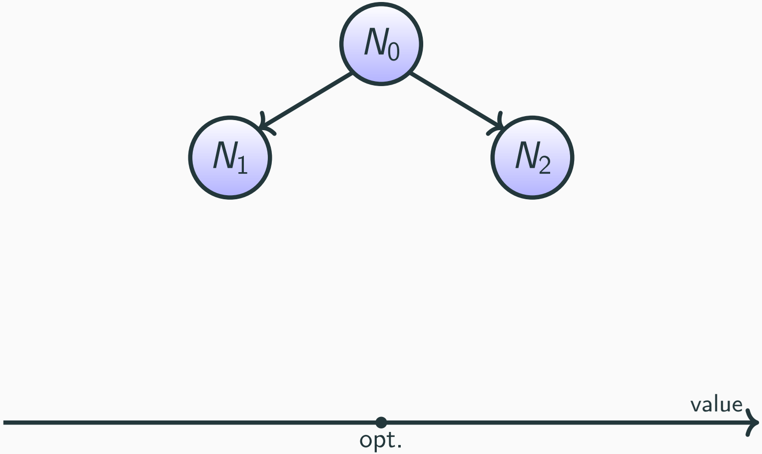
Tree exploration

N_0

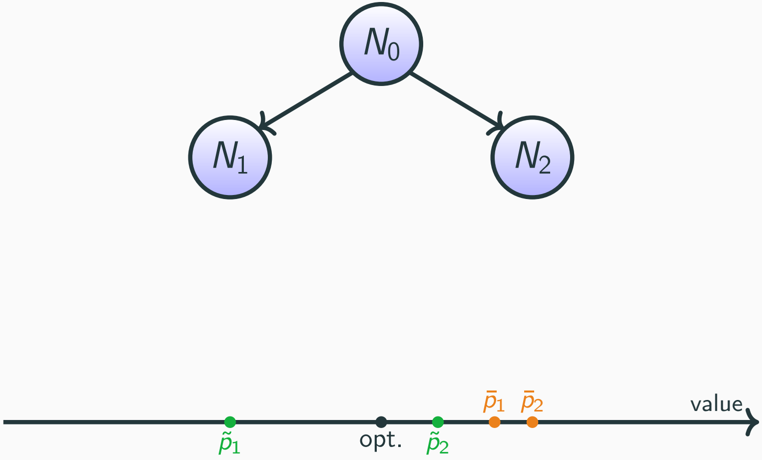
Tree exploration



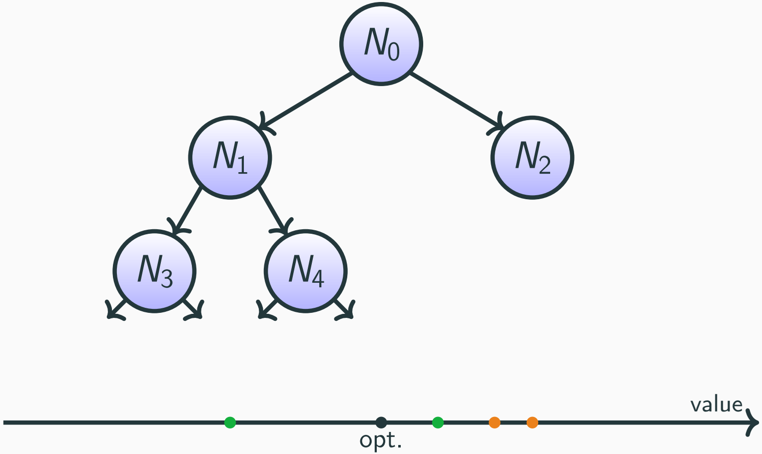
Tree exploration



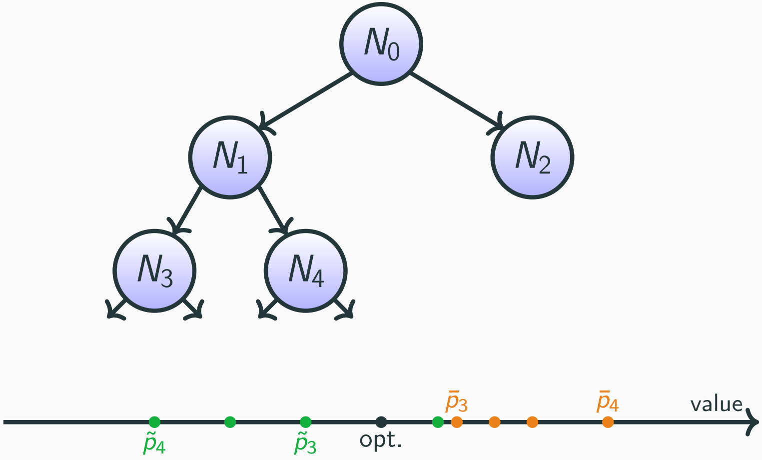
Tree exploration



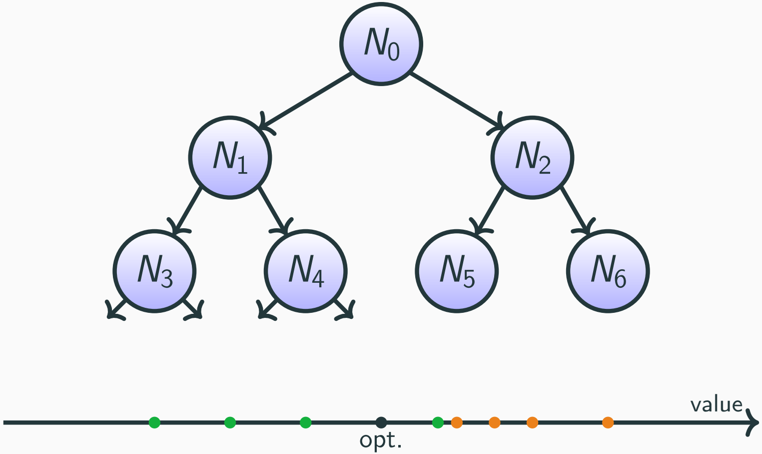
Tree exploration



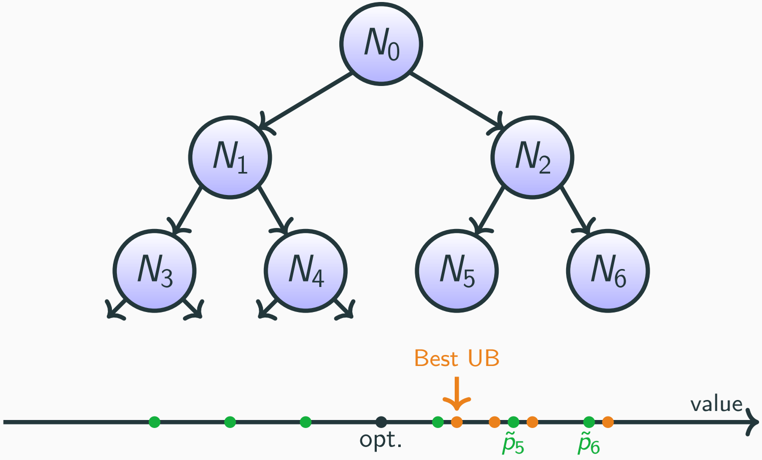
Tree exploration



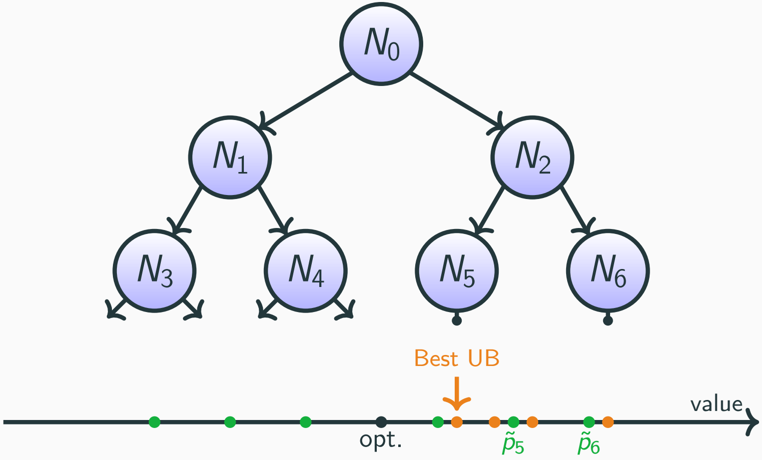
Tree exploration



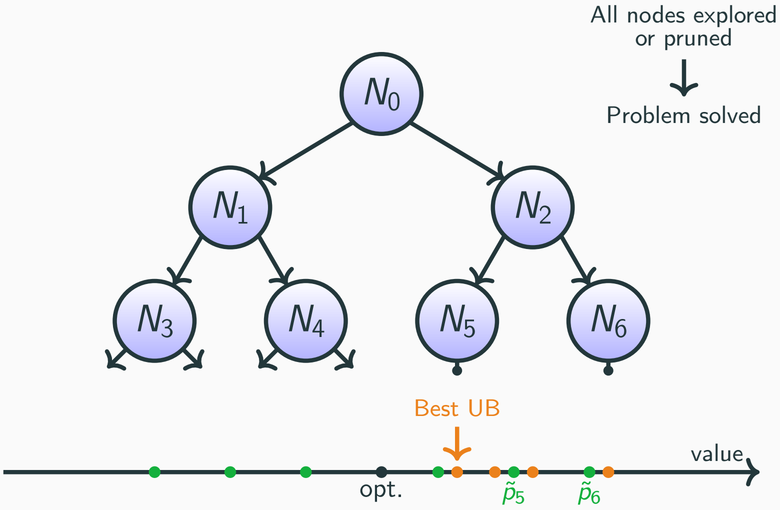
Tree exploration



Tree exploration



Tree exploration



Ingredients

ℓ_0-penalized problem

$$\min_{\mathrm{x}} \; F(\mathrm{x}) + \lambda\|\mathrm{x}\|_0 + G(\mathrm{x})$$

What do we need

- A branching strategy
- A bounding strategy
- Something generic with respect to $F(\cdot)$ and $G(\cdot)$

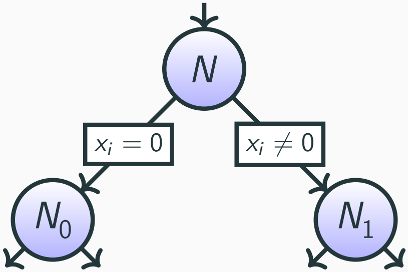
Generic solution method

Branching

Force the (non-)nullity of an entry

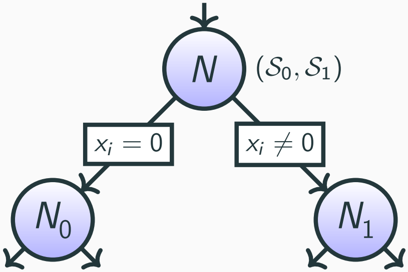
Branching

Force the (non-)nullity of an entry



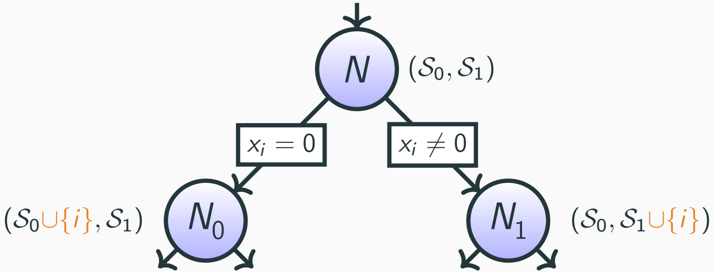
Branching

Force the (non-)nullity of an entry



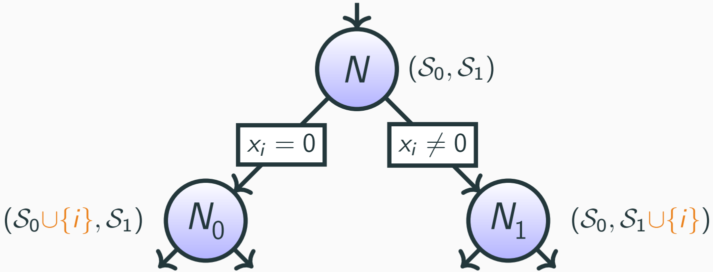
Branching

Force the (non-)nullity of an entry



Branching

Force the (non-)nullity of an entry



Node problem

$$\begin{cases} \min_{\mathbf{x}} & F(\mathbf{x}) + \lambda \|\mathbf{x}\|_0 + G(\mathbf{x}) \\ \text{s.t.} & \mathbf{x}_{\mathcal{S}_0} = 0,\ \mathbf{x}_{\mathcal{S}_1} \neq 0 \end{cases}$$

Upper bounding

N_k

$(\mathcal{S}_0, \mathcal{S}_1)$

Node problem

$$\begin{cases} \min_{\mathbf{x}} & F(\mathbf{x}) + \lambda\|\mathbf{x}\|_0 + G(\mathbf{x}) \\ \text{s.t.} & \mathbf{x}_{\mathcal{S}_0} = 0,\ \mathbf{x}_{\mathcal{S}_1} \neq 0 \end{cases}$$

Still NP-hard unless all entries are fixed

N_k

$(\mathcal{S}_0, \mathcal{S}_1)$

Node problem

$$\begin{cases} \min_{\mathbf{x}} & F(\mathbf{x}) + \lambda\|\mathbf{x}\|_0 + G(\mathbf{x}) \\ \text{s.t.} & \mathbf{x}_{\mathcal{S}_0} = 0,\ \mathbf{x}_{\mathcal{S}_1} \neq 0 \end{cases}$$

Still NP-hard unless all entries are fixed

$\boxed{\text{Fix } x_i = 0 \text{ for all } i \notin \mathcal{S}_1}$

Upper bounding

$(\mathcal{S}_0, \mathcal{S}_1)$

Node problem

$$\begin{cases} \min_{\mathrm{x}} & F(\mathrm{x}) + \lambda\|\mathrm{x}\|_0 + G(\mathrm{x}) \\ \text{s.t.} & \mathrm{x}_{\mathcal{S}_0} = 0,\ \mathrm{x}_{\mathcal{S}_1} \neq 0 \end{cases}$$

Still NP-hard unless all entries are fixed

$\downarrow$

$\boxed{\text{Fix } x_i = 0 \text{ for all } i \notin \mathcal{S}_1}$

$\downarrow$

Upper bounding problem

$$\min_{\mathrm{x}}\ F(\mathrm{x}_{\mathcal{S}_1}) + \lambda|\mathcal{S}_1| + G(\mathrm{x}_{\mathcal{S}_1})$$

Upper bounding

$(\mathcal{S}_0, \mathcal{S}_1)$

Node problem

$$\begin{cases} \min_{\mathrm{x}} & F(\mathrm{x}) + \lambda\|\mathrm{x}\|_0 + G(\mathrm{x}) \\ \text{s.t.} & \mathrm{x}_{\mathcal{S}_0} = 0,\ \mathrm{x}_{\mathcal{S}_1} \neq 0 \end{cases}$$

Still NP-hard unless all entries are fixed

$\boxed{\text{Fix } x_i = 0 \text{ for all } i \notin \mathcal{S}_1}$

Upper bounding problem

$$\min_{\mathrm{x}}\ F(\mathrm{x}_{\mathcal{S}_1}) + \lambda|\mathcal{S}_1| + G(\mathrm{x}_{\mathcal{S}_1})$$

Convex problem

Upper bound of good quality

Node problem

$$\begin{cases} \min_{\mathbf{x}} & F(\mathbf{x}) + \lambda\|\mathbf{x}\|_0 + G(\mathbf{x}) \\ \text{s.t.} & \mathbf{x}_{\mathcal{S}_0} = 0,\ \mathbf{x}_{\mathcal{S}_1} \neq 0 \end{cases}$$

Lower bounding

N_k

$(\mathcal{S}_0, \mathcal{S}_1)$

Node problem

$$\begin{cases} \min_{\mathbf{x}} & F(\mathbf{x}) + \lambda \|\mathbf{x}\|_0 + G(\mathbf{x}) \\ \text{s.t.} & \mathbf{x}_{\mathcal{S}_0} = 0, \ \mathbf{x}_{\mathcal{S}_1} \neq 0 \end{cases}$$

$\min_{\mathbf{x}} P(\mathbf{x})$

Lower bounding

N_k

$(\mathcal{S}_0, \mathcal{S}_1)$

Node problem

$$\begin{cases} \min_{\mathrm{x}} & F(\mathrm{x}) + \lambda\|\mathrm{x}\|_0 + G(\mathrm{x}) \\ \text{s.t.} & \mathrm{x}_{\mathcal{S}_0} = 0,\ \mathrm{x}_{\mathcal{S}_1} \neq 0 \end{cases}$$

$\min_{\mathrm{x}} P(\mathrm{x})$

Use bi-conjugacy $P^{\star\star}(\mathrm{x}) \leq P(\mathrm{x})$

Lower bounding

N_k

$(\mathcal{S}_0, \mathcal{S}_1)$

Node problem

$$\begin{cases} \min_{\mathrm{x}} & F(\mathrm{x}) + \lambda\|\mathrm{x}\|_0 + G(\mathrm{x}) \\ \text{s.t.} & \mathrm{x}_{\mathcal{S}_0} = 0,\ \mathrm{x}_{\mathcal{S}_1} \neq 0 \end{cases}$$

$\min_{\mathrm{x}} P(\mathrm{x})$

Use bi-conjugacy $P^{\star\star}(\mathrm{x}) \leq P(\mathrm{x})$

Lower bounding problem

$\min_{\mathrm{x}} P^{\star\star}(\mathrm{x})$

Lower bounding

N_k

$(\mathcal{S}_0, \mathcal{S}_1)$

Node problem

$$\begin{cases} \min_{\mathrm{x}} & F(\mathrm{x}) + \lambda\|\mathrm{x}\|_0 + G(\mathrm{x}) \\ \text{s.t.} & \mathrm{x}_{\mathcal{S}_0} = 0,\ \mathrm{x}_{\mathcal{S}_1} \neq 0 \end{cases}$$

$\min_{\mathrm{x}} P(\mathrm{x})$

Use bi-conjugacy $P^{\star\star}(\mathrm{x}) \leq P(\mathrm{x})$

Lower bounding problem

$\min_{\mathrm{x}} P^{\star\star}(\mathrm{x})$

Bi-conjugate computation is also NP-hard

😵

Lower bounding

Our solution: Compute the bi-conjugate of a part of the problem

Node problem

$$\begin{cases} \min_{\mathbf{x}} & F(\mathbf{x}) + \lambda\|\mathbf{x}\|_0 + G(\mathbf{x}) \\ \text{s.t.} & \mathbf{x}_{\mathcal{S}_0} = 0,\ \mathbf{x}_{\mathcal{S}_1} \neq 0 \end{cases}$$

Our solution: Compute the bi-conjugate of a part of the problem

Node problem

$$\begin{cases} \min_x & F(x) + \lambda\|x\|_0 + G(x) \\ \text{s.t.} & x_{\mathcal{S}_0} = 0,\ x_{\mathcal{S}_1} \neq 0 \end{cases}$$

$$\Big\downarrow \boxed{\text{Reformulation}}$$

$$\min_x \ F(x) + H(x)$$

Lower bounding

Our solution: Compute the bi-conjugate of a part of the problem

Node problem

$$\begin{cases} \min_x & F(x) + \lambda\|x\|_0 + G(x) \\ \text{s.t.} & x_{\mathcal{S}_0} = 0,\ x_{\mathcal{S}_1} \neq 0 \end{cases}$$

$\downarrow$ Reformulation

$$\min_x \ F(x) + H(x)$$

$\downarrow$ Bi-conjugacy

$$\min_x \ F(x) + H^{\star\star}(x)$$

Lower-bounding problem

Lower bounding

$$H(\mathrm{x}) = \lambda\|\mathrm{x}\|_{\mathbf{0}} + G(\mathrm{x}) + \mathbb{I}(\mathrm{x}_{\mathcal{S}_{\mathbf{0}}} = 0) + \mathbb{I}(\mathrm{x}_{\mathcal{S}_{\mathbf{1}}} \neq 0)$$

Lower bounding

$$H(\mathrm{x}) = \lambda\|\mathrm{x}\|_0 + G(\mathrm{x}) + \mathbb{I}(\mathrm{x}_{\mathcal{S}_0} = 0) + \mathbb{I}(\mathrm{x}_{\mathcal{S}_1} \neq 0)$$

$$\Big\downarrow \ \boxed{\text{Separable}}$$

$$H^{\star\star}(\mathrm{x}) = \textstyle\sum_i H_i^{\star\star}(x_i)$$

Lower bounding

$$H(\mathrm{x}) = \lambda\|\mathrm{x}\|_0 + G(\mathrm{x}) + \mathbb{I}(\mathrm{x}_{\mathcal{S}_0} = 0) + \mathbb{I}(\mathrm{x}_{\mathcal{S}_1} \neq 0)$$

Separable

$$H^{\star\star}(\mathrm{x}) = \textstyle\sum_i H_i^{\star\star}(x_i)$$

$i \in \mathcal{S}_0$

$$H_i(x_i) = \mathbb{I}(x_i = 0)$$

Already convex: $H_i^{\star\star}(\cdot) = H_i(\cdot)$

Lower bounding

$$H(\mathrm{x}) = \lambda\|\mathrm{x}\|_0 + G(\mathrm{x}) + \mathbb{I}(\mathrm{x}_{\mathcal{S}_0} = 0) + \mathbb{I}(\mathrm{x}_{\mathcal{S}_1} \neq 0)$$

Separable

$$H^{\star\star}(\mathrm{x}) = \textstyle\sum_i H_i^{\star\star}(x_i)$$

$i \in \mathcal{S}_0$

$$H_i(x_i) = \mathbb{I}(x_i = 0)$$

Already convex: $H_i^{\star\star}(\cdot) = H_i(\cdot)$

$i \in \mathcal{S}_1$

$$H_i(x_i) = \lambda + G_i(x_i)$$

Already convex: $H_i^{\star\star}(\cdot) = H_i(\cdot)$

Lower bounding

$$H(\mathrm{x}) = \lambda\|\mathrm{x}\|_0 + G(\mathrm{x}) + \mathbb{I}(\mathrm{x}_{\mathcal{S}_0} = 0) + \mathbb{I}(\mathrm{x}_{\mathcal{S}_1} \neq 0)$$

Separable

$$H^{\star\star}(\mathrm{x}) = \textstyle\sum_i H_i^{\star\star}(x_i)$$

$i \in \mathcal{S}_0$

$$H_i(x_i) = \mathbb{I}(x_i = 0)$$

Already convex: $H_i^{\star\star}(\cdot) = H_i(\cdot)$

i unfixed

$$H_i(x_i) = \lambda\|x_i\|_0 + G_i(x_i)$$

$i \in \mathcal{S}_1$

$$H_i(x_i) = \lambda + G_i(x_i)$$

Already convex: $H_i^{\star\star}(\cdot) = H_i(\cdot)$

Lower bounding

$$H(\mathrm{x}) = \lambda\|\mathrm{x}\|_0 + G(\mathrm{x}) + \mathbb{I}(\mathrm{x}_{\mathcal{S}_0} = 0) + \mathbb{I}(\mathrm{x}_{\mathcal{S}_1} \neq 0)$$

Separable

$$H^{\star\star}(\mathrm{x}) = \textstyle\sum_i H_i^{\star\star}(x_i)$$

$i \in \mathcal{S}_0$

$$H_i(x_i) = \mathbb{I}(x_i = 0)$$

Already convex: $H_i^{\star\star}(\cdot) = H_i(\cdot)$

i unfixed

$$H_i(x_i) = \lambda\|x_i\|_0 + G_i(x_i)$$

Closed form bi-conjugate

$$H_i^{\star\star}(x_i) = \begin{cases} \tau|x_i| & \text{if } |x_i| \leq \mu \\ \lambda + G_i(x_i) & \text{otherwise} \end{cases}$$

$i \in \mathcal{S}_1$

$$H_i(x_i) = \lambda + G_i(x_i)$$

Already convex: $H_i^{\star\star}(\cdot) = H_i(\cdot)$

Lower bounding

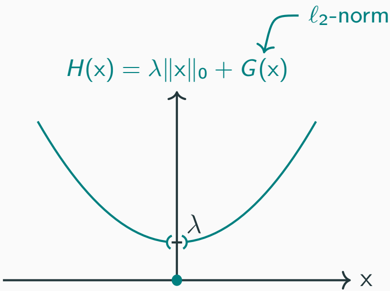
Graphical intuition

$$H_i^{\star\star}(x_i) = \begin{cases} \tau|x_i| & \text{if } |x_i| \leq \mu \\ \lambda + G_i(x_i) & \text{otherwise} \end{cases}$$

Lower bounding

Graphical intuition

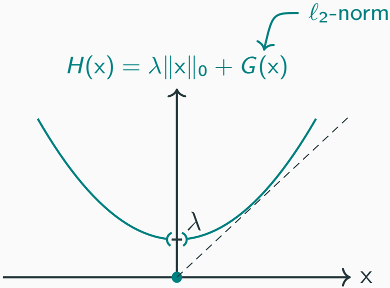
$$H_i^{\star\star}(x_i) = \begin{cases} \tau|x_i| & \text{if } |x_i| \leq \mu \\ \lambda + G_i(x_i) & \text{otherwise} \end{cases}$$



Lower bounding

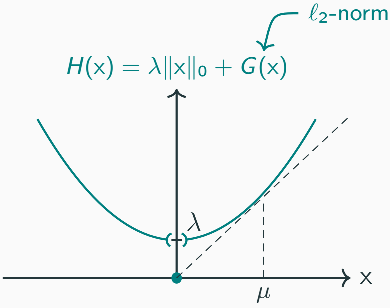
Graphical intuition

$$H_i^{\star\star}(x_i) = \begin{cases} \tau|x_i| & \text{if } |x_i| \leq \mu \\ \lambda + G_i(x_i) & \text{otherwise} \end{cases}$$



Graphical intuition

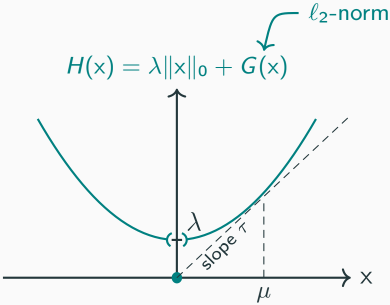
$$H_i^{\star\star}(x_i) = \begin{cases} \tau|x_i| & \text{if } |x_i| \leq \mu \\ \lambda + G_i(x_i) & \text{otherwise} \end{cases}$$



Lower bounding

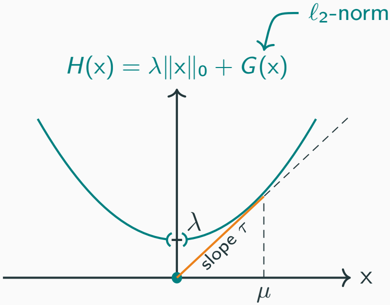
Graphical intuition

$$H_i^{\star\star}(x_i) = \begin{cases} \tau|x_i| & \text{if } |x_i| \leq \mu \\ \lambda + G_i(x_i) & \text{otherwise} \end{cases}$$



Graphical intuition

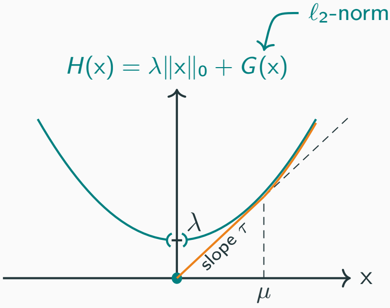
$$H_i^{\star\star}(x_i) = \begin{cases} \tau|x_i| & \text{if } |x_i| \leq \mu \\ \lambda + G_i(x_i) & \text{otherwise} \end{cases}$$



Lower bounding

Graphical intuition

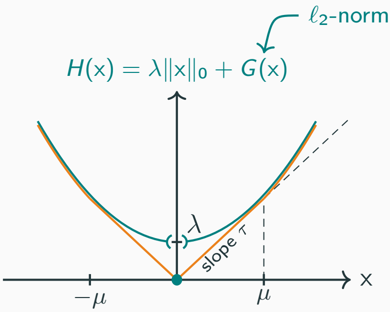
$$H_i^{\star\star}(x_i) = \begin{cases} \tau|x_i| & \text{if } |x_i| \leq \mu \\ \lambda + G_i(x_i) & \text{otherwise} \end{cases}$$



Lower bounding

Graphical intuition

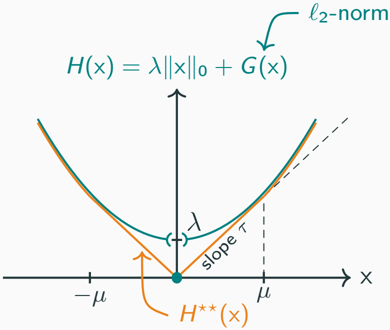
$$H_i^{\star\star}(x_i) = \begin{cases} \tau|x_i| & \text{if } |x_i| \leq \mu \\ \lambda + G_i(x_i) & \text{otherwise} \end{cases}$$



Lower bounding

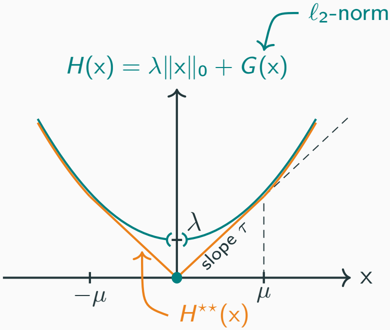
Graphical intuition

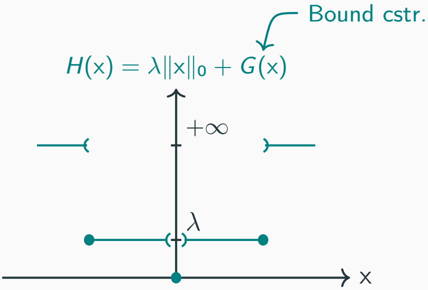
$$H_i^{\star\star}(x_i) = \begin{cases} \tau|x_i| & \text{if } |x_i| \leq \mu \\ \lambda + G_i(x_i) & \text{otherwise} \end{cases}$$



Graphical intuition

$$H_i^{\star\star}(x_i) = \begin{cases} \tau|x_i| & \text{if } |x_i| \leq \mu \\ \lambda + G_i(x_i) & \text{otherwise} \end{cases}$$

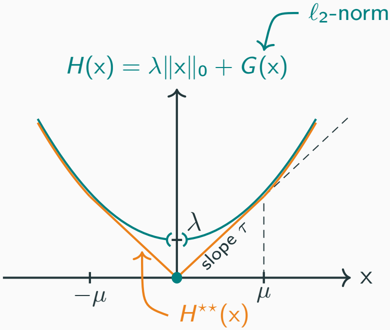


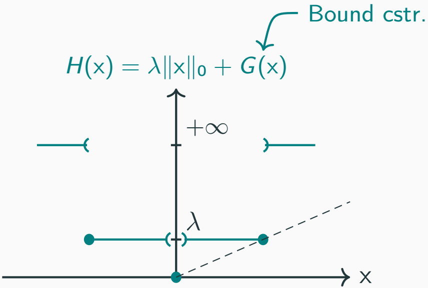


Lower bounding

Graphical intuition

$$H_i^{\star\star}(x_i) = \begin{cases} \tau|x_i| & \text{if } |x_i| \leq \mu \\ \lambda + G_i(x_i) & \text{otherwise} \end{cases}$$

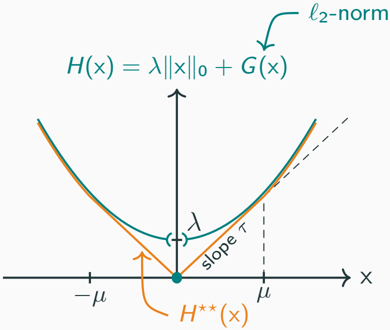


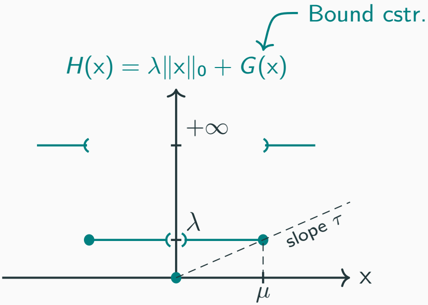


Lower bounding

Graphical intuition

$$H_i^{\star\star}(x_i) = \begin{cases} \tau|x_i| & \text{if } |x_i| \leq \mu \\ \lambda + G_i(x_i) & \text{otherwise} \end{cases}$$

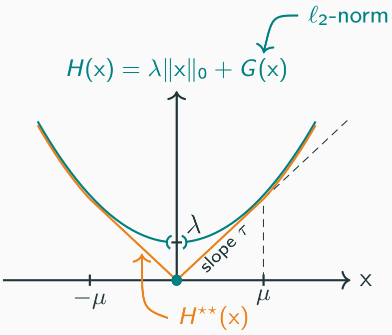


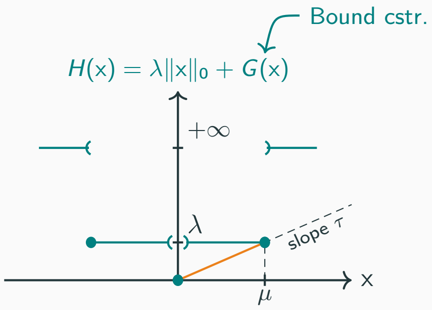


Lower bounding

Graphical intuition

$$H_i^{\star\star}(x_i) = \begin{cases} \tau|x_i| & \text{if } |x_i| \leq \mu \\ \lambda + G_i(x_i) & \text{otherwise} \end{cases}$$

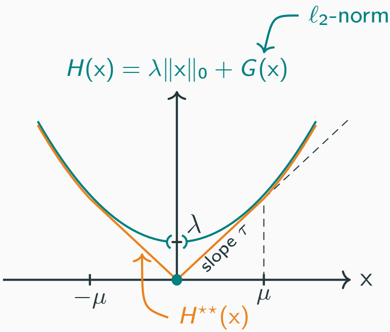


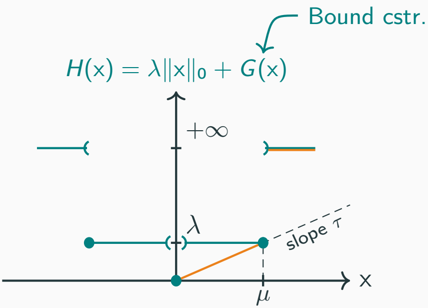


Lower bounding

Graphical intuition

$$H_i^{\star\star}(x_i) = \begin{cases} \tau|x_i| & \text{if } |x_i| \leq \mu \\ \lambda + G_i(x_i) & \text{otherwise} \end{cases}$$

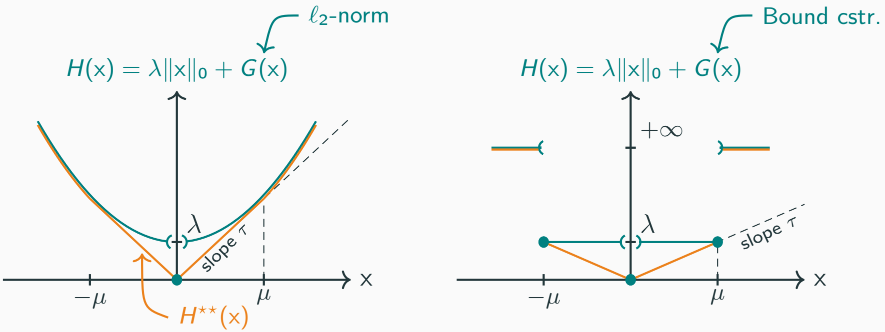




Lower bounding

Graphical intuition

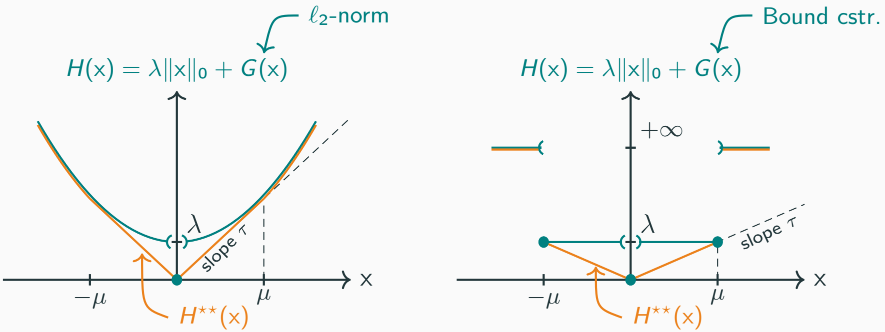
$$H_i^{\star\star}(x_i) = \begin{cases} \tau|x_i| & \text{if } |x_i| \leq \mu \\ \lambda + G_i(x_i) & \text{otherwise} \end{cases}$$



Lower bounding

Graphical intuition

$$H_i^{\star\star}(x_i) = \begin{cases} \tau|x_i| & \text{if } |x_i| \leq \mu \\ \lambda + G_i(x_i) & \text{otherwise} \end{cases}$$



Let's sum up !

ℓ_0-penalized problem

$$\min_x \; F(x) + \lambda\|x\|_0 + G(x)$$

- Branch-and-Bound algorithm

Let's sum up !

ℓ_0-penalized problem

$$\min_{x} \; F(x) + \lambda \|x\|_0 + G(x)$$

- Branch-and-Bound algorithm
- Branching strategy
  - Fix the nullity of entries

Let's sum up !

ℓ_0-penalized problem

$$\min_x \; F(x) + \lambda\|x\|_0 + G(x)$$

- Branch-and-Bound algorithm
- Branching strategy
  - Fix the nullity of entries
- Upper bounding problem
  - Set all the unfixed entries to zero
  - Convex problem

Let's sum up !

ℓ_0-penalized problem

$$\min_x \; F(x) + \lambda \|x\|_0 + G(x)$$

- Branch-and-Bound algorithm
- Branching strategy
  - Fix the nullity of entries
- Upper bounding problem
  - Set all the unfixed entries to zero
  - Convex problem
- Lower bounding problem
  - Bi-conjugate of a part of the objective value
  - Closed-form expression
  - Convex problem

Numerical results

Fast

Generic

TheoGuyard/**El0ps.jl**

An Exact L0-penalized Problem Solver

Allows user-defined $F(\cdot)$ and $G(\cdot)$

The secret sauce

Branch-and-Bound

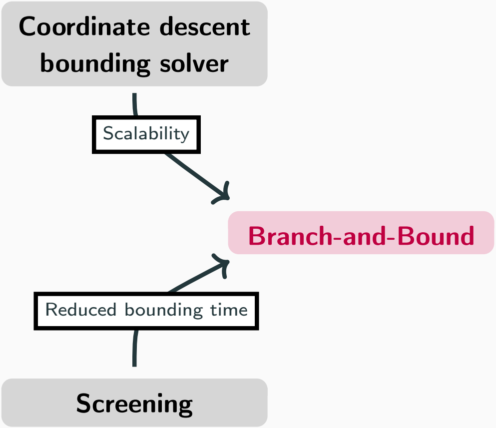
The secret sauce

Coordinate descent bounding solver

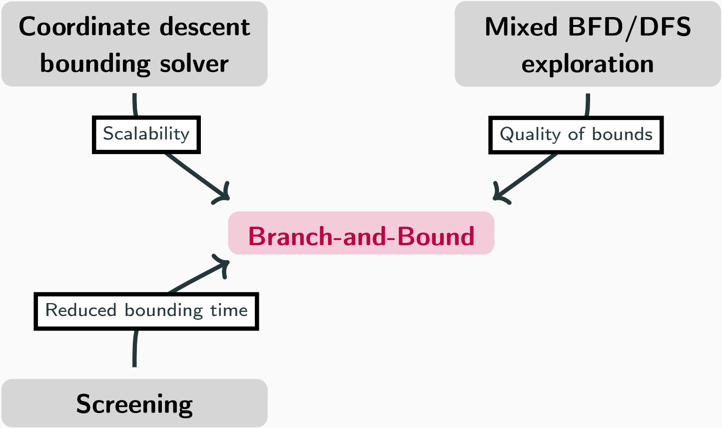
Scalability

Branch-and-Bound

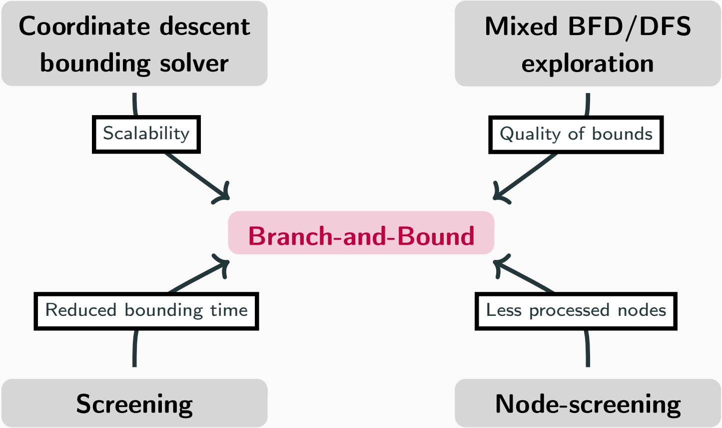
The secret sauce



The secret sauce



The secret sauce



Numerical results

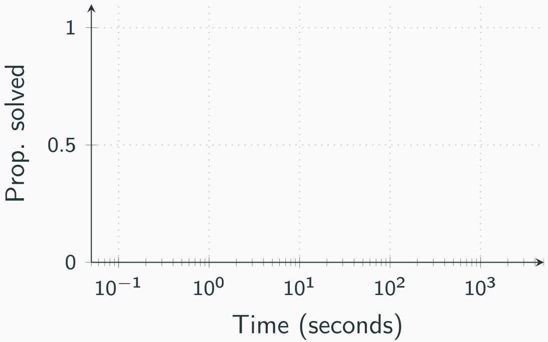
$$\min_x \ F(x) + \lambda\|x\|_0 + G(x)$$

Dataset : Sparse regression
F(·) : Least-squares loss
G(·) : Bound constraints
$\boldsymbol{\lambda}$: Set statistically

Numerical results

$$\min_x \ F(x) + \lambda \|x\|_0 + G(x)$$

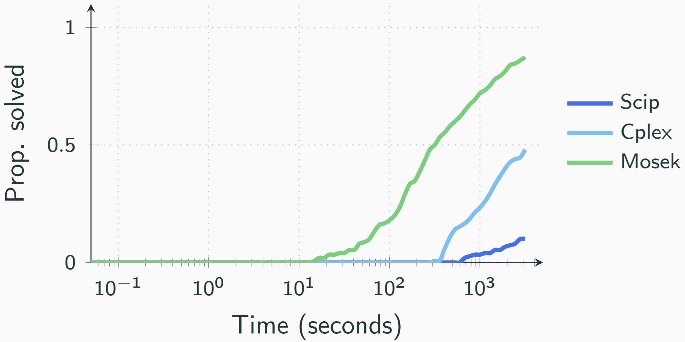
Dataset : Sparse regression
F(·) : Least-squares loss
G(·) : Bound constraints
λ : Set statistically



Numerical results

$$\min_x \; F(x) + \lambda \|x\|_0 + G(x)$$

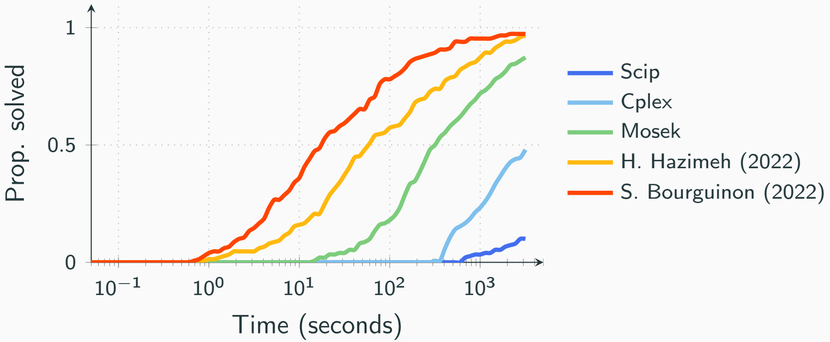
Dataset : Sparse regression
F(·) : Least-squares loss
G(·) : Bound constraints
λ : Set statistically



Numerical results

$$\min_x \; F(x) + \lambda\|x\|_0 + G(x)$$

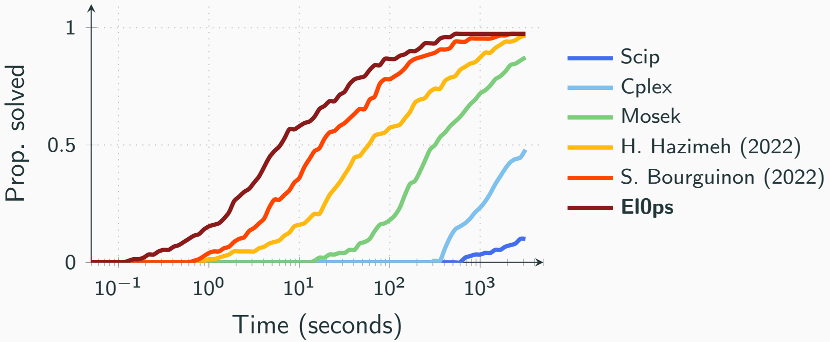
Dataset : Sparse regression
F(·) : Least-squares loss
G(·) : Bound constraints
$\boldsymbol{\lambda}$: Set statistically



Numerical results

$$\min_x \; F(x) + \lambda\|x\|_0 + G(x)$$

Dataset : Sparse regression
F(·) : Least-squares loss
G(·) : Bound constraints
λ : Set statistically



Numerical results

$$\min_{x} \; F(x) + \lambda \|x\|_0 + G(x)$$

Dataset : Sparse regression

F(·) : Least-squares loss

G(·) : **ℓ_2-norm**

λ : Set statistically

Numerical results

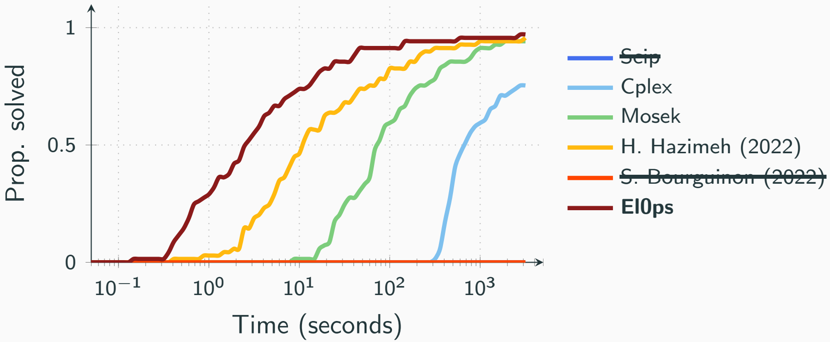
$$\min_x \ F(x) + \lambda \|x\|_0 + G(x)$$

Dataset : Sparse regression

$\mathbf{F(\cdot)}$: Least-squares loss

$\mathbf{G(\cdot)}$: **ℓ_2-norm**

$\boldsymbol{\lambda}$: Set statistically



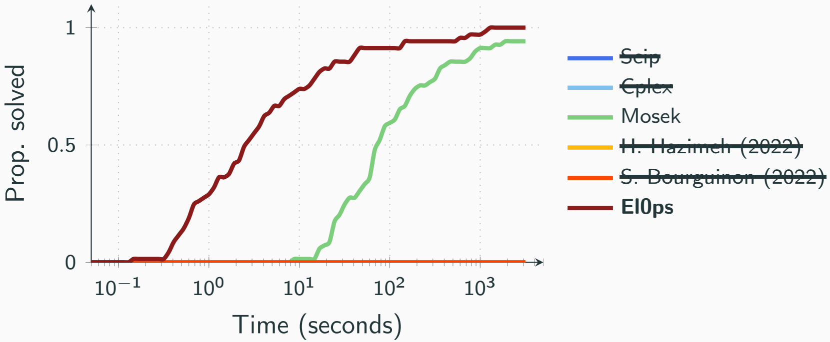
Numerical results

$$\min_x \; F(x) + \lambda \|x\|_0 + G(x)$$

Dataset	:	Sparse **classification**
F(·)	:	**Logistic loss**
G(·)	:	Bound constraints
$\boldsymbol{\lambda}$	:	Set statistically

$$\min_x \; F(x) + \lambda \|x\|_0 + G(x)$$

Dataset : Sparse **classification**
F(·) : **Logistic loss**
G(·) : Bound constraints
λ : Set statistically



Conclusion

ℓ_0-penalized problem

$$\min_x \; F(x) + \lambda \|x\|_0 + G(x)$$

ℓ_0-penalized problem

$$\min_x \; F(x) + \lambda\|x\|_0 + G(x)$$

- Generic solution method
  - $F(\cdot)$ convex
  - $G(\cdot)$ convex and separable

Conclusion

ℓ_0-penalized problem

$$\min_{x} \; F(x) + \lambda \|x\|_0 + G(x)$$

- Generic solution method
  - $F(\cdot)$ convex
  - $G(\cdot)$ convex and separable
- Significant gains against competitors
  - Factor $\times 10^6$ against off-the-shelf solvers
  - Factor $\times 10^2$ against specialized solvers

Conclusion

ℓ_0-penalized problem

$$\min_x \; F(x) + \lambda \|x\|_0 + G(x)$$

- Generic solution method
  - $F(\cdot)$ convex
  - $G(\cdot)$ convex and separable
- Significant gains against competitors
  - Factor $\times 10^6$ against off-the-shelf solvers
  - Factor $\times 10^2$ against specialized solvers
- We do not trade efficiency for flexibility
  - Handle broader practical cases
  - Address more challenging instances

Question time

TheoGuyard/El0ps.jl

An Exact L0-penalized Problem Solver