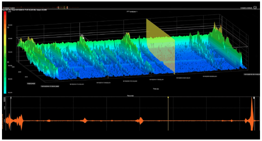
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Institute for Operations Research
Department of Mathematics

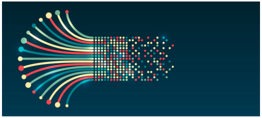
Constrained Sparse Approximation Over the Cube

Sabrina Bruckmeier Christoph Hunkenschröder Robert Weismantel

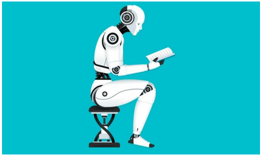
Sparse Approximation is the problem of identifying a subset that most accurately models an observation.
(Sebastian Ament and Carla Gomes)



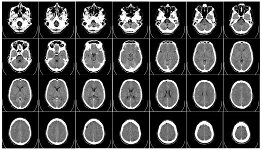
(Signal Recovery)



(Pattern Recognition)



(Machine Learning)



(Computed Tomography)

(Portfolio Selection)

Mathematical Description.

- Given:
  - Matrix $A \in \mathbb{Z}^{m \times n}$,
  - Vector $b \in \mathbb{Z}^m$,
  - Integer $\sigma \in [n]$.
- Task:

$$\min\{\|b - Ax\|_2 : x \in [0,1]^n, \|x\|_0 \leq \sigma\}. \qquad (P_0)$$

- In General: NP-Hard [1]
- Idea: Let's fix m.

[1] Tropp, J.A. (2004). "Greed is good: Algorithmic results for sparse approximation". IEEE Transactions on Information Theory. 50 (10): 2231–2242.

Mathematical Description.

- Given:
  - Matrix $A \in \mathbb{Z}^{m \times n}$,
  - Vector $b \in \mathbb{Z}^{m}$,
  - Integer $\sigma \in [n]$.
- Task:

$$\min\{\|b - Ax\|_2 : x \in [0,1]^n, \|x\|_0 \leq \sigma\}. \qquad (P_0)$$

- In General: NP-Hard [1]
- Idea: Let's fix m.

[1] Tropp, J.A. (2004). "Greed is good: Algorithmic results for sparse approximation". IEEE Transactions on Information Theory. 50 (10): 2231–2242.

Mathematical Description.

- Given:
  - Matrix $A \in \mathbb{Z}^{m \times n}$,
  - Vector $b \in \mathbb{Z}^m$,
  - Integer $\sigma \in [n]$.
- Task:

$$\min\{\|b - Ax\|_2 : x \in [0,1]^n, \|x\|_0 \leq \sigma\}. \qquad (P_0)$$

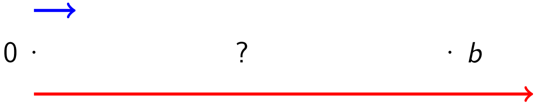
- In General: NP-Hard [1]
- Idea: Let's fix m.

[1] Tropp, J.A. (2004). "Greed is good: Algorithmic results for sparse approximation". IEEE Transactions on Information Theory. 50 (10): 2231–2242.

That's easy!

$$\min\{\|b - Ax\|_2 : x \in [0,1]^n, \|x\|_0 \leq \sigma\}.$$

- $m = 1$



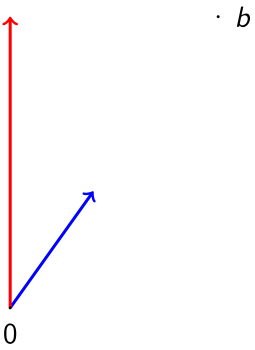
That's easy!

$$\min\{\|b - Ax\|_2 : x \in [0,1]^n, \|x\|_0 \leq \sigma\}.$$

- $m = 1$

Or not?

- $m = 2$

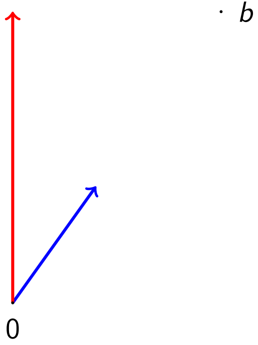


Theorem

(P_0) is NP-hard, even if $m = 2$.

Or not?

- $m = 2$



Theorem

(P_0) *is NP-hard, even if* $m = 2$.

Assumptions

- $\|A\|_\infty$ bounded
- m bounded
- $m \ll n$

$$A = \begin{bmatrix} \star & \star & \star & \star & \dots & \star & \star & \star \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ \star & \star & \star & \star & \dots & \star & \star & \star \end{bmatrix}$$

ℓ_1-Relaxation.

$$\min\{\|b - Ax\|_2 : x \in [0,1]^n, \|x\|_0 \le \sigma\} \qquad (P_0)$$

$$\min\{\|b - Ax\|_2 : x \in [0,1]^n, \|x\|_1 \le \sigma\} \qquad (P_1)$$

So far

Algorithms

- need A to satisfy (hard to verify) properties.
- depend highly on the input.
- succeed only with a certain probability.

Our contribution

Independent of A we give

- Probabilistic analysis for random targets b.
- Proximity result between (P_0) and (P_1).
- Deterministic algorithm polynomial in n provided m and $\|A\|_\infty$ constant.

So far

Algorithms

- need A to satisfy (hard to verify) properties.
- depend highly on the input.
- succeed only with a certain probability.

Our contribution

Independent of A we give

- Probabilistic analysis for random targets b.
- Proximity result between (P_0) and (P_1).
- Deterministic algorithm polynomial in n provided m and $\|A\|_\infty$ constant.

So far

Algorithms

- need A to satisfy (hard to verify) properties.
- depend highly on the input.
- succeed only with a certain probability.

Our contribution

Independent of A we give

- Probabilistic analysis for random targets b.
- Proximity result between (P_0) and (P_1).
- Deterministic algorithm polynomial in n provided m and $\|A\|_\infty$ constant.

So far

Algorithms

- need A to satisfy (hard to verify) properties.
- depend highly on the input.
- succeed only with a certain probability.

Our contribution

Independent of A we give

- Probabilistic analysis for random targets b.
- Proximity result between (P_0) and (P_1).
- Deterministic algorithm polynomial in n provided m and $\|A\|_\infty$ constant.

So far

Algorithms

- need A to satisfy (hard to verify) properties.
- depend highly on the input.
- succeed only with a certain probability.

Our contribution

Independent of A we give

- Probabilistic analysis for random targets b.
- Proximity result between (P_0) and (P_1).
- Deterministic algorithm polynomial in n provided m and $\|A\|_\infty$ constant.

So far

Algorithms

- need A to satisfy (hard to verify) properties.
- depend highly on the input.
- succeed only with a certain probability.

Our contribution

Independent of A we give

- Probabilistic analysis for random targets b.
- Proximity result between (P_0) and (P_1).
- Deterministic algorithm polynomial in n provided m and $\|A\|_\infty$ constant.

Few fractional entries for (P_0).

- Feasible vector x with support S.
- Set of (P_0)-feasible points with same objective value:

$$P_S(x) := \{y \in \mathbb{R}^{|S|} : Ax = A_S y, 0 \leq y \leq 1\}.$$

- $P_S(x)$:
  - polyhedron,
  - non-empty,
  - has at least one vertex v.
- At v at least $|S| - m$ inequalities of the form $0 \leq y \leq 1$ are tight.

Few fractional entries for (P_0).

- Feasible vector x with support S.
- Set of (P_0)-feasible points with same objective value:

$$P_S(x) := \{y \in \mathbb{R}^{|S|} : Ax = A_S y, 0 \leq y \leq 1\}.$$

- $P_S(x)$:
  - polyhedron,
  - non-empty,
  - has at least one vertex v.
- At v at least $|S| - m$ inequalities of the form $0 \leq y \leq 1$ are tight.

Few fractional entries for (P_0).

- Feasible vector x with support S.
- Set of (P_0)-feasible points with same objective value:

$$P_S(x) := \{y \in \mathbb{R}^{|S|} : Ax = A_S y, 0 \le y \le 1\}.$$

- $P_S(x)$:
  - polyhedron,
  - non-empty,
  - has at least one vertex v.
- At v at least $|S| - m$ inequalities of the form $0 \le y \le 1$ are tight.

Few fractional entries for (P_0).

- Feasible vector x with support S.
- Set of (P_0)-feasible points with same objective value:

$$P_S(x) := \{y \in \mathbb{R}^{|S|} : Ax = A_S y, 0 \leq y \leq 1\}.$$

- $P_S(x)$:
  - polyhedron,
  - non-empty,
  - has at least one vertex v.
- At v at least $|S| - m$ inequalities of the form $0 \leq y \leq 1$ are tight.

Few fractional entries for (P_0).

- Feasible vector x with support S.
- Set of (P_0)-feasible points with same objective value:

$$P_S(x) := \{y \in \mathbb{R}^{|S|} : Ax = A_S y, 0 \leq y \leq 1\}.$$

- $P_S(x)$:
  - polyhedron,
  - non-empty,
  - has at least one vertex v.
- At v at least $|S| - m$ inequalities of the form $0 \leq y \leq 1$ are tight.

Few fractional entries for (P_0).

Lemma

There exists a solution of (P_0) that has at most m fractional entries.

Few fractional entries for (P_1).

Lemma

There exists a solution of (P_1) that has at most m fractional entries.

Section 1

The ℓ_1-relaxation for Random Targets b.

Which vectors b are "easy" target vectors?

- Set of points representable with the ℓ_1-relaxation

$$Q := \{Ax \in \mathbb{R}^m : x \in [0,1]^n, \|x\|_1 \leq \sigma\}.$$

- If b is "deep" inside Q, then (P_0) is easy.

Theorem

If $b \in \frac{\sigma-m+1}{\sigma} Q$, then an optimal solution of (P_1) solves (P_0).

Which vectors b are "easy" target vectors?

- Set of points representable with the ℓ_1-relaxation

$$Q := \{Ax \in \mathbb{R}^m : x \in [0,1]^n, \|x\|_1 \leq \sigma\}.$$

- If b is "deep" inside Q, then (P_0) is easy.

Theorem

If $b \in \frac{\sigma - m + 1}{\sigma} Q$, then an optimal solution of (P_1) solves (P_0).

Which vectors b are "easy" target vectors?

- Set of points representable with the ℓ_1-relaxation

$$Q := \{Ax \in \mathbb{R}^m : x \in [0,1]^n, \|x\|_1 \leq \sigma\}.$$

- If b is "deep" inside Q, then (P_0) is easy.

Theorem

If $b \in \frac{\sigma - m + 1}{\sigma} Q$, then an optimal solution of (P_1) solves (P_0).

Which vectors b are "easy" target vectors?

- Set of points representable with the ℓ_1-relaxation

$$Q := \{Ax \in \mathbb{R}^m : x \in [0,1]^n, \|x\|_1 \leq \sigma\}.$$

- If b is "deep" inside Q, then (P_0) is easy.

Theorem

If $b \in \frac{\sigma - m + 1}{\sigma} Q$, then an optimal solution of (P_1) solves (P_0).

Which vectors b are "easy" target vectors?

- $Q := \{Ax \in \mathbb{R}^m : x \in [0,1]^n, \|x\|_1 \leq \sigma\}$.

Theorem

If $b \in \frac{\sigma - m + 1}{\sigma} Q$, then an optimal solution of (P_1) solves (P_0).

Proof (Sketch).

- Polyhedron $\{x \in [0,1]^n : Ax = b, \|x\|_1 \leq \sigma - m + 1\}$.
- Vertex v has at most $\sigma - m + 1$ integral non-zero entries.
- v has at most m fractional entries.
- $\|v\|_0 \leq \sigma$.
- v is optimal for (P_0).

□

Which vectors b are "easy" target vectors?

- $Q := \{Ax \in \mathbb{R}^m : x \in [0,1]^n, \|x\|_1 \le \sigma\}$.

Theorem

If $b \in \frac{\sigma - m + 1}{\sigma} Q$, then an optimal solution of (P_1) solves (P_0).

Proof (Sketch).

- Polyhedron $\{x \in [0,1]^n : Ax = b, \|x\|_1 \le \sigma - m + 1\}$.
- Vertex v has at most $\sigma - m + 1$ integral non-zero entries.
- v has at most m fractional entries.
- $\|v\|_0 \le \sigma$.
- v is optimal for (P_0).

□

Which vectors b are "easy" target vectors?

- $Q := \{Ax \in \mathbb{R}^m : x \in [0,1]^n, \|x\|_1 \leq \sigma\}$.

Theorem

If $b \in \frac{\sigma - m + 1}{\sigma} Q$, then an optimal solution of (P_1) solves (P_0).

Proof (Sketch).

- Polyhedron $\{x \in [0,1]^n : Ax = b, \|x\|_1 \leq \sigma - m + 1\}$.
- Vertex v has at most $\sigma - m + 1$ integral non-zero entries.
- v has at most m fractional entries.
- $\|v\|_0 \leq \sigma$.
- v is optimal for (P_0).

Which vectors b are "easy" target vectors?

- $Q := \{Ax \in \mathbb{R}^m : x \in [0,1]^n, \|x\|_1 \leq \sigma\}$.

Theorem

If $b \in \frac{\sigma - m + 1}{\sigma} Q$, then an optimal solution of (P_1) solves (P_0).

Proof (Sketch).

- Polyhedron $\{x \in [0,1]^n : Ax = b, \|x\|_1 \leq \sigma - m + 1\}$.
- Vertex v has at most $\sigma - m + 1$ integral non-zero entries.
- v has at most m fractional entries.
- $\|v\|_0 \leq \sigma$.
- v is optimal for (P_0).

□

Which vectors b are "easy" target vectors?

- $Q := \{Ax \in \mathbb{R}^m : x \in [0,1]^n, \|x\|_1 \leq \sigma\}$.

Theorem

If $b \in \frac{\sigma - m + 1}{\sigma} Q$, then an optimal solution of (P_1) solves (P_0).

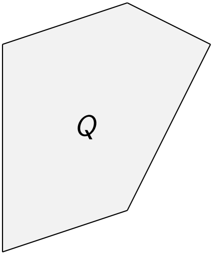
Proof (Sketch).

- Polyhedron $\{x \in [0,1]^n : Ax = b, \|x\|_1 \leq \sigma - m + 1\}$.
- Vertex v has at most $\sigma - m + 1$ integral non-zero entries.
- v has at most m fractional entries.
- $\|v\|_0 \leq \sigma$.
- v is optimal for (P_0).

□

Which vectors b are "easy" target vectors?

- $Q := \{Ax \in \mathbb{R}^m : x \in [0,1]^n, \|x\|_1 \leq \sigma\}$.



Which vectors b are "easy" target vectors?

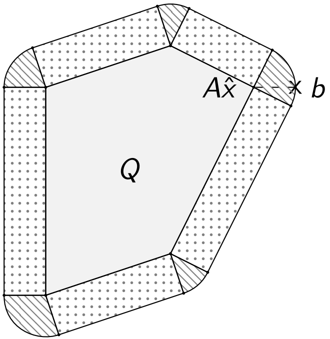


Figure: The sampling of the vector b from $Q + \lambda B$

Which vectors b are "easy" target vectors?

- If b is far outside of Q, then (P_0) is easy with high probability.

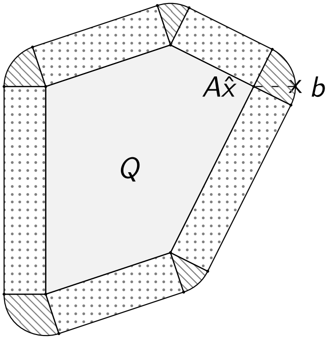


Figure: The sampling of the vector b from $Q + \lambda B$

Which vectors b are "easy" target vectors?

Theorem

If b is sampled uniformly at random from the convex set $Q + \lambda B$, then with probability at least

$$\rho = \left(\frac{\lambda}{\lambda + \sigma m \|A\|_\infty} \right)^m$$

there exists an optimal solution of (P_1) that is optimal for (P_0).

- Example: If $\lambda = 2m^2\sigma\|A\|_\infty$, then $\rho \geq \frac{1}{2}$ by Bernoulli's inequality.

Which vectors b are "easy" target vectors?

- Conversely, if b is close to the boundary of Q, then the probability that an optimal solution of (P_1) solves (P_0) is almost 0.

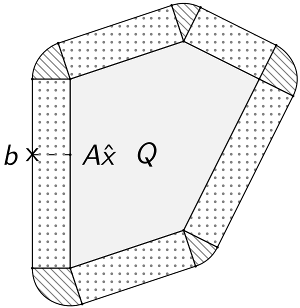


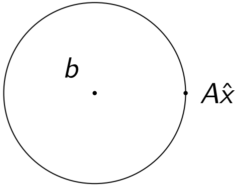
Figure: The sampling of the vector b from $Q + \lambda B$

Section 2

Proximity between (P_1) and (P_0).

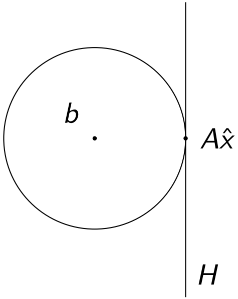
Separation Lemma.

- $\hat{x}$ optimal solution of (P_1).



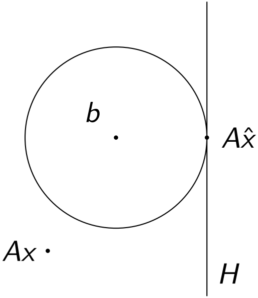
Separation Lemma.

- $\hat{x}$ optimal solution of (P_1).

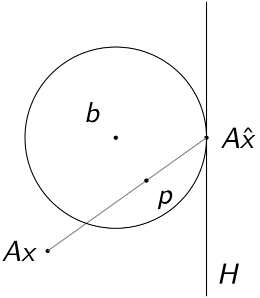


Separation Lemma.

- $\hat{x}$ optimal solution of (P_1).
- x feasible point of (P_0).



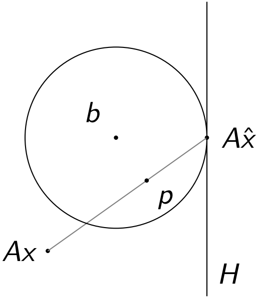
Separation Lemma.



Lemma

H separates b from all vectors Ax with x feasible for (P_0).

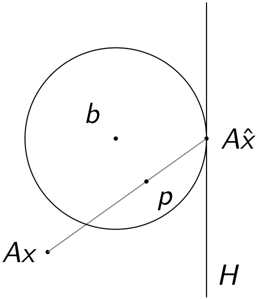
Separation Lemma.



Lemma

H separates b from all vectors Ax with x feasible for (P_0).

Perturbation Lemma.



Lemma

If we perturb $\hat{x}$ along the fractional entries, we will remain in H.

Proximity Result.

Theorem

For an optimal solution $x^\star$ to (P_0) we have

$$\|Ax^\star - A\hat{x}\|_2 \leq 2m^{3/2}\|A\|_\infty.$$

Proximity Result.

Theorem

For an optimal solution $x^\star$ to (P_0) we have

$$\|Ax^\star - A\hat{x}\|_2 \leq 2m^{3/2}\|A\|_\infty.$$

Proof (Sketch)

- Reduce support of $\hat{x}$ by "filling up" the fractional entries $\rightarrow y$.
- y is feasible for (P_0).
- By the Perturbation Lemma $Ay \in H$.
- Separation Lemma + Standard Linear Algebra yields the result.

Proximity Result.

Theorem

For an optimal solution $x^\star$ to (P_0) we have

$$\|Ax^\star - A\hat{x}\|_2 \leq 2m^{3/2}\|A\|_\infty.$$

Proof (Sketch)

- Reduce support of $\hat{x}$ by "filling up" the fractional entries $\rightarrow y$.
- y is feasible for (P_0).
- By the Perturbation Lemma $Ay \in H$.
- Separation Lemma + Standard Linear Algebra yields the result.

Proximity Result.

Theorem

For an optimal solution $x^\star$ to (P_0) we have

$$\|Ax^\star - A\hat{x}\|_2 \leq 2m^{3/2}\|A\|_\infty.$$

Proof (Sketch)

- Reduce support of $\hat{x}$ by "filling up" the fractional entries $\rightarrow y$.
- y is feasible for (P_0).
- By the Perturbation Lemma $Ay \in H$.
- Separation Lemma + Standard Linear Algebra yields the result.

Proximity Result.

Theorem

For an optimal solution $x^\star$ to (P_0) we have

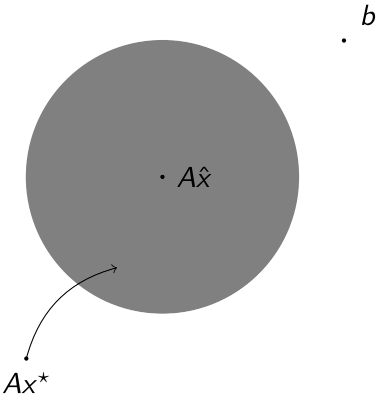
$$\|Ax^\star - A\hat{x}\|_2 \leq 2m^{3/2}\|A\|_\infty.$$

Proof (Sketch)

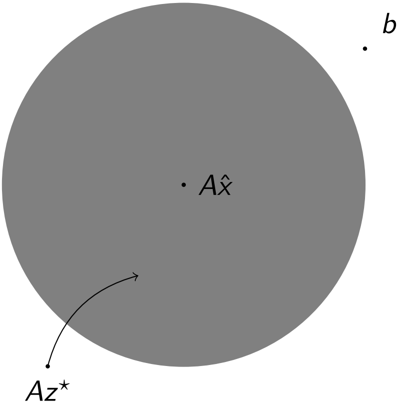
- Reduce support of $\hat{x}$ by "filling up" the fractional entries $\rightarrow y$.
- y is feasible for (P_0).
- By the Perturbation Lemma $Ay \in H$.
- Separation Lemma + Standard Linear Algebra yields the result.

Section 3

A Deterministic Algorithm.

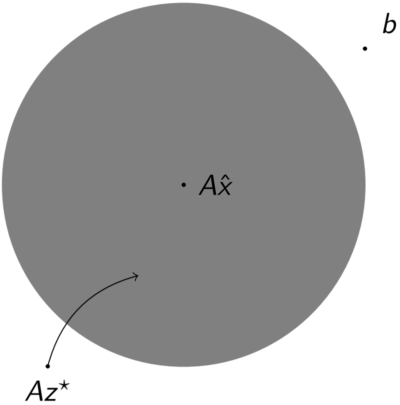

b
$A\hat{x}$
$Ax^\star$

Main Idea: Decompose $x^\star = z^\star + f^\star$.



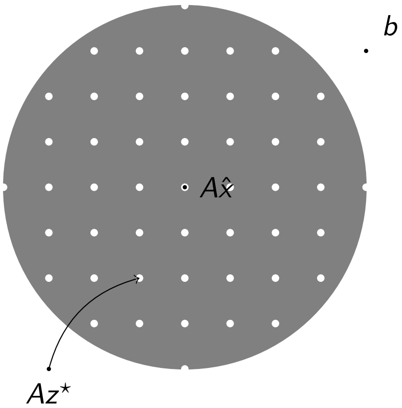
1. Guess support $\text{supp}(f^\star)$ of fractional entries.

Main Idea: Decompose $x^\star = z^\star + f^\star$.



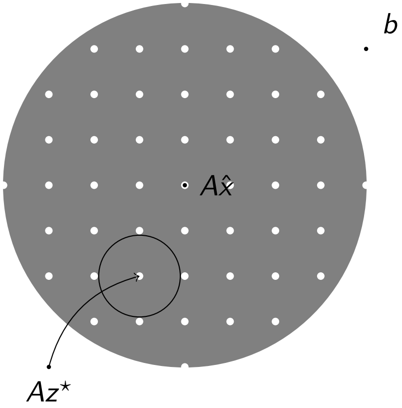
1. Guess support $\mathrm{supp}(f^\star)$ of fractional entries.

Main Idea: Decompose $x^\star = z^\star + f^\star$.



2. Establish candidate set $Z^\star$ for $z^\star$.

Main Idea: Decompose $x^\star = z^\star + f^\star$.



3. Match each $z \in Z^\star$ with its fractional part f.

Arithmetic Cost

- Guess support $\text{supp}(f^\star)$ of fractional entries.
  - $\text{supp}(f^\star) \leq m$.
  - Minimal index set of fractional entries uses distinct columns of A.
  - There are at most $(2\|A\|_\infty + 1)^m$ distinct columns.

Lemma

There are at most $(2\|A\|_\infty + 1)^{m^2}$ potentially different index sets $\text{supp}(f^\star)$.

Arithmetic Cost

- Guess support $\text{supp}(f^\star)$ of fractional entries.
  - $\text{supp}(f^\star) \leq m$.
  - Minimal index set of fractional entries uses distinct columns of A.
  - There are at most $(2\|A\|_\infty + 1)^m$ distinct columns.

Lemma

There are at most $(2\|A\|_\infty + 1)^{m^2}$ potentially different index sets $\text{supp}(f^\star)$.

Arithmetic Cost

- Guess support supp($f^\star$) of fractional entries.
  - supp($f^\star$) $\leq m$.
  - Minimal index set of fractional entries uses distinct columns of A.
  - There are at most $(2\|A\|_\infty + 1)^m$ distinct columns.

Lemma

There are at most $(2\|A\|_\infty + 1)^{m^2}$ potentially different index sets supp($f^\star$).

Arithmetic Cost

- Guess support $\mathrm{supp}(f^\star)$ of fractional entries.
  - $\mathrm{supp}(f^\star) \leq m$.
  - Minimal index set of fractional entries uses distinct columns of A.
  - There are at most $(2\|A\|_\infty + 1)^m$ distinct columns.

Lemma

There are at most $(2\|A\|_\infty + 1)^{m^2}$ potentially different index sets $\mathrm{supp}(f^\star)$.

Arithmetic Cost

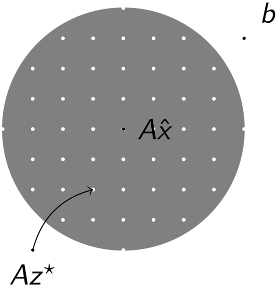
- Guess support supp($f^\star$) of fractional entries.
  - supp($f^\star$) $\leq m$.
  - Minimal index set of fractional entries uses distinct columns of A.
  - There are at most $(2\|A\|_\infty + 1)^m$ distinct columns.

Lemma

There are at most $(2\|A\|_\infty + 1)^{m^2}$ potentially different index sets supp($f^\star$).

Arithmetic Cost

- Establish candidate set $Z^\star$ for $z^\star$.



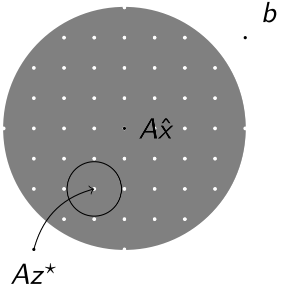
Theorem

Compute $Z^\star$ by solving at most $\mathcal{O}(m^{\frac{3}{2}}\|A\|_\infty)^m$ LIPs

$$A_{\setminus f^\star} y = b',$$
$$\sum_{i=1}^{n-m} y_i \leq \sigma - m,$$
$$y \in \{0,1\}^{n-m}.$$

Arithmetic Cost

- Match each $z \in Z^\star$ with its fractional part f.



Theorem

A solution of

$$\min\{\|(b - A_{\setminus f^\star} z^\star) - A_{f^\star} g\|_2 : \; g \in [0,1]^m\}$$

can be computed in $\mathcal{O}(3^m m^3)$ *arithmetic operations.*

Arithmetic Cost.

Theorem

The arithmetic cost of finding an optimal solution to (P_0) is

$$(m\|A\|_\infty)^{\mathcal{O}(m^2)} \cdot poly(n, \ln(\|b\|_1)).$$