Hardness and Algorithms for Sparse Optimization

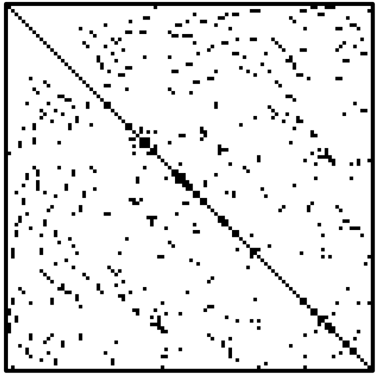
Eric Price
UT Austin

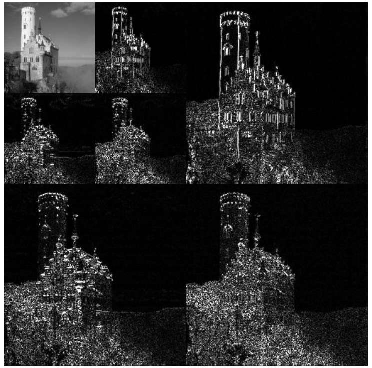
Sandeep Silwal
MIT

Samson Zhou
CMU

What is this paper about?

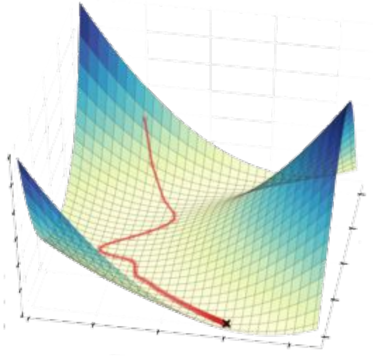
- ➢ Sparsity in data is a natural pattern
- ➢ Study optimization problems with sparsity constraints

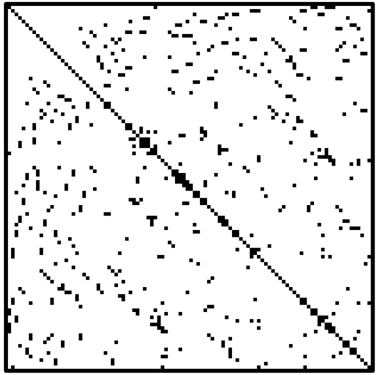


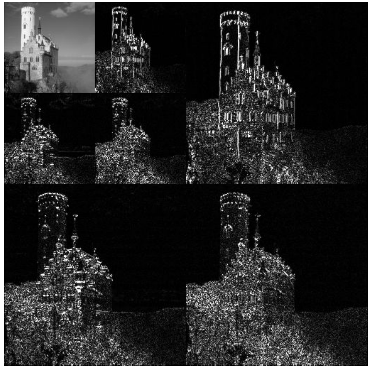


What is this paper about?

- ➢ Sparsity in data is a natural pattern
- ➢ Study optimization problems with sparsity constraints







What is the paper about?

- ➢ Lower bounds and reductions between natural sparse optimization problems
- ➢ Unified approach to upper bound algorithms for sparse optimization problems

What is the paper about?

- Lower bounds and reductions between natural sparse optimization problems
- Unified approach to upper bound algorithms for sparse optimization problems

Two heroes of the story

- **Sparse regression:** "Find a k-sparse solution to $Ax = b$"

Two heroes of the story

➢ **Sparse regression:** "Find a k-sparse solution to $Ax = b$"

Given A in $R^{n \times d}$, b in R^n, and integer k, find x with $\|x\|_0 \leq k$ to minimize $\|Ax-b\|_2$

Two heroes of the story

- ➢ **Sparse regression:** "Find a k-sparse solution to Ax = b"

Given A in $R^{n \times d}$, b in R^n, and integer k, find x with $\|x\|_0 \le k$ to minimize $\|Ax-b\|_2$

- ➢ **Robust regression:** "Choose k constraints to ignore in Ax = b"

Two heroes of the story

➢ **Sparse regression:** "Find a k-sparse solution to $Ax = b$"

Given A in $R^{n \times d}$, b in R^n, and integer k, find x with $\|x\|_0 \leq k$ to minimize $\|Ax-b\|_2$

➢ **Robust regression:** "Choose k constraints to ignore in $Ax = b$"

Find subset T of size $\leq k$ and x to minimize $\| (Ax-b)_T \|_2$ where $(Ax-b)_T$ zeros out all coordinates in T

Two heroes of the story

➢ **Sparse regression:** "Find a k-sparse solution to Ax = b"

Given A in $R^{n \times d}$, b in R^n, and integer k, find x with $\|x\|_0 \leq k$ to minimize $\|Ax-b\|_2$

➢ **Robust regression:** "Choose k constraints to ignore in Ax = b"

Find subset T of size ≤ k and x to minimize $\| (Ax-b)_T \|_2$ where $(Ax-b)_T$ zeros out all coordinates in T

Very related problems: We show robust regression reduces to sparse regression!

A common algorithmic approach

We give a common algorithmic approach for sparse and robust regression (and other problems)

Based on “Meet in the Middle”

A common algorithmic approach

We give a common algorithmic approach for sparse and robust regression (and other problems)

Based on "Meet in the Middle"

- ➢ Consider sparse regression: "Find a k-sparse x to solve Ax = b"
- ➢ Suppose there exists y with $\|y\|_0 = k$ with Ay = b (**Realizable case**)

A common algorithmic approach

We give a common algorithmic approach for sparse and robust regression (and other problems)

Based on “Meet in the Middle”

- ➢ Consider sparse regression: “Find a k-sparse x to solve Ax = b”
- ➢ Suppose there exists y with $\|y\|_0 = k$ with Ay = b (**Realizable case**)

Brute force algorithm: Guess support of y: roughly **d^k time** where A in $R^{n \times d}$

Prior work: work very hard to get roughly ~ **d^{k-1} time** (Har-Peled, Indyk, Mahabadi ‘18)

A common algorithmic approach

We give a common algorithmic approach for sparse and robust regression (and other problems)

Based on “Meet in the Middle”

- ➢ Consider sparse regression: “Find a k-sparse x to solve Ax = b”
- ➢ Suppose there exists y with $\|y\|_0 = k$ with Ay = b (**Realizable case**)

Brute force algorithm: Guess support of y: roughly **d^k time** where A in $R^{n \times d}$

Prior work: work very hard to get roughly ~ **d^{k-1} time** (Har-Peled, Indyk, Mahabadi ‘18)

Our result: ~ **$d^{k/2}$** time.

A common algorithmic approach

We give a common algorithmic approach for sparse and robust regression (and other problems)

Based on "Meet in the Middle"

- ➢ Consider sparse regression: "Find a k-sparse x to solve Ax = b"
- ➢ Suppose there exists y with $\|y\|_0 = k$ with Ay = b (**Realizable case**)

(1) Enumerate over all k/2 sparse z and compute Az

(2) Check if there exists k/2 sparse z' such that Az' = b-Az

(3) Output z'+z as solution

Only ~ $\mathbf{d^{k/2}}$ cost!

A common algorithmic approach

We give a common algorithmic approach for sparse and robust regression (and other problems)

Based on “Meet in the Middle”

- ➢ Consider sparse regression: “Find a k-sparse x to solve Ax = b”
- ➢ Suppose there exists y with $\|y\|_0 = k$ with Ay = b (**Realizable case**)

(1) Enumerate over all k/2 sparse z and compute Az ← Enumerate over net

(2) Check if there exists k/2 sparse z’ such that Az’ = b-Az

(3) Output z’+z as solution

Only ~ $\mathbf{d^{k/2}}$ cost!

A common algorithmic approach

We give a common algorithmic approach for sparse and robust regression (and other problems)

Based on "Meet in the Middle"

- ➢ Consider sparse regression: "Find a k-sparse x to solve Ax = b"
- ➢ Suppose there exists y with $\|y\|_0 = k$ with Ay = b (**Realizable case**)

(1) Enumerate over all k/2 sparse z and compute Az ← Enumerate over net

(2) Check if there exists k/2 sparse z' such that Az' = b-Az ← Check for z' using nearest neighbor data structure on {Az}

(3) Output z'+z as solution

Only ~ $\mathbf{d^{k/2}}$ cost!

A common algorithmic approach

We give a common algorithmic approach for sparse and robust regression (and other problems)

Theorem 1.4. *Given* $A \in \mathbb{R}^{n \times d}$ *and* $b \in \mathbb{R}^n$ *such that there exists a* k*-sparse vector* x *satisfying* $Ax = b$*, there exists an algorithm that returns a* k*-sparse vector* $z \in \mathbb{R}^d$ *satisfying* $\|Az - b\|_2 \le \varepsilon$ *in time*

$$\min_{c \ge 1} O\left(nk \cdot \left(\frac{12c \cdot d \cdot \|b\|_2}{\varepsilon} \right)^{\frac{k}{2} \cdot (1 + 1/(2c^2 - 1))} \right).$$

A common algorithmic approach

We give a common algorithmic approach for sparse and robust regression (and other problems)

Theorem 1.4. *Given $A \in \mathbb{R}^{n \times d}$ and $b \in \mathbb{R}^n$ such that there exists a k-sparse vector x satisfying $Ax = b$, there exists an algorithm that returns a k-sparse vector $z \in \mathbb{R}^d$ satisfying $\|Az - b\|_2 \leq \varepsilon$ in time*

$$\min_{c \geq 1} O\left(nk \cdot \left(\frac{12c \cdot d \cdot \|b\|_2}{\varepsilon}\right)^{\frac{k}{2} \cdot (1 + 1/(2c^2 - 1))}\right).$$

Similar improvement for robust regression ($\mathbf{n^k \rightarrow n^{k/2}}$)

Other results

Upper bounds:

- Sparse PCA problem: “Find a k sparse unit vector u to maximize $u^T A u$”

Lower bounds side:

- Bicriteria hardness for sparse regression: “Problem still hard even if we relax sparsity and approximation constraints”
- Fine grained hardness for robust regression: Evidence of $\mathbf{n^{k/2}}$ hardness for robust regression

Thank you!

Please see our paper for full details!