COMPRESSED SENSING

Hu Fu, Oct 2023

SHUFE

SPARSE SIGNALS

SPARSE SIGNALS

- High-dimensional but sparse signals arise in many applications

SPARSE SIGNALS

- High-dimensional but sparse signals arise in many applications
  - Image/media files are often sparse when expressed w.r.t. the right bases

SPARSE SIGNALS

- High-dimensional but sparse signals arise in many applications
  - Image/media files are often sparse when expressed w.r.t. the right bases
    - E.g. wavelet transform

SPARSE SIGNALS

- High-dimensional but sparse signals arise in many applications
  - Image/media files are often sparse when expressed w.r.t. the right bases
    - E.g. wavelet transform
    - $\mathbf{x} \in \mathbb{R}^n$, with $\|\mathbf{x}\|_0 \leq s$. We say $\mathbf{x}$ is s-sparse

SPARSE SIGNALS

- High-dimensional but sparse signals arise in many applications
  - Image/media files are often sparse when expressed w.r.t. the right bases
    - E.g. wavelet transform
    - $\mathbf{x} \in \mathbb{R}^n$, with $\|\mathbf{x}\|_0 \leq s$. We say $\mathbf{x}$ is s-sparse
- We often have linear measurements of such signals

SPARSE SIGNALS

- High-dimensional but sparse signals arise in many applications
  - Image/media files are often sparse when expressed w.r.t. the right bases
    - E.g. wavelet transform
    - $\mathbf{x} \in \mathbb{R}^n$, with $\|\mathbf{x}\|_0 \leq s$. We say $\mathbf{x}$ is s-sparse
- We often have linear measurements of such signals
  - $\mathbf{y} = A\mathbf{x}$, where A is a matrix in $\mathbb{R}^{m \times n}$, with $m \ll n$

SPARSE SIGNALS

- High-dimensional but sparse signals arise in many applications
  - Image/media files are often sparse when expressed w.r.t. the right bases
    - E.g. wavelet transform
    - $\mathbf{x} \in \mathbb{R}^n$, with $\|\mathbf{x}\|_0 \leq s$. We say $\mathbf{x}$ is s-sparse
- We often have linear measurements of such signals
  - $\mathbf{y} = A\mathbf{x}$, where A is a matrix in $\mathbb{R}^{m \times n}$, with $m \ll n$
  - Given $\mathbf{y}$, can we recover $\mathbf{x}$? We can design both the measurements A and the recovery algorithm.

SPARSE RECOVERY WITH COUNT SKETCH

SPARSE RECOVERY WITH COUNT SKETCH

- Recall that with Count Sketch, we were able to recover sparse signals

SPARSE RECOVERY WITH COUNT SKETCH

- Recall that with Count Sketch, we were able to recover sparse signals
- By taking the largest (in absolute value) s coordinates of the sketch, with high probability, we get an s-sparse $\tilde{\mathbf{x}}$ s.t. $\|\tilde{\mathbf{x}} - \mathbf{x}\|_2 \leq (1 + \epsilon) E_2^s(\mathbf{x})$

SPARSE RECOVERY WITH COUNT SKETCH

- Recall that with Count Sketch, we were able to recover sparse signals
- By taking the largest (in absolute value) s coordinates of the sketch, with high probability, we get an s-sparse $\tilde{\mathbf{x}}$ s.t. $\|\tilde{\mathbf{x}} - \mathbf{x}\|_2 \leq (1+\epsilon)E_2^s(\mathbf{x})$
  - $E_2^s(\mathbf{x})$ is the ℓ_2-norm of $\mathbf{x}$ with its largest k coordinates zeroed out

SPARSE RECOVERY WITH COUNT SKETCH

- Recall that with Count Sketch, we were able to recover sparse signals
- By taking the largest (in absolute value) s coordinates of the sketch, with high probability, we get an s-sparse $\tilde{\mathbf{x}}$ s.t. $\|\tilde{\mathbf{x}} - \mathbf{x}\|_2 \leq (1+\epsilon)E_2^s(\mathbf{x})$
  - $E_2^s(\mathbf{x})$ is the ℓ_2-norm of $\mathbf{x}$ with its largest k coordinates zeroed out
  - If $\mathbf{x}$ is s-sparse, with high probability $\tilde{\mathbf{x}}$ is an exact recovery

SPARSE RECOVERY WITH COUNT SKETCH

- Recall that with Count Sketch, we were able to recover sparse signals
- By taking the largest (in absolute value) s coordinates of the sketch, with high probability, we get an s-sparse $\tilde{\mathbf{x}}$ s.t. $\|\tilde{\mathbf{x}} - \mathbf{x}\|_2 \leq (1+\epsilon)E_2^s(\mathbf{x})$
    - $E_2^s(\mathbf{x})$ is the ℓ_2-norm of $\mathbf{x}$ with its largest k coordinates zeroed out
    - If $\mathbf{x}$ is s-sparse, with high probability $\tilde{\mathbf{x}}$ is an exact recovery
    - Count Sketch consists of randomized linear measurements of $\mathbf{x}$. $\tilde{\mathbf{x}}$ is computed from them

NON-UNIFORM VS. UNIFORM RECOVERY

NON-UNIFORM VS. UNIFORM RECOVERY

- Non-uniform schemes: $\forall \mathbf{x} \in \mathbb{R}^n, \ \Pr[\tilde{\mathbf{x}} \text{ recovers } \mathbf{x}] \geq 1 - \delta$

NON-UNIFORM VS. UNIFORM RECOVERY

- Non-uniform schemes: $\forall \mathbf{x} \in \mathbb{R}^n$, $\Pr[\tilde{\mathbf{x}} \text{ recovers } \mathbf{x}] \geq 1 - \delta$
  - Count Sketch gives us such a guarantee

NON-UNIFORM VS. UNIFORM RECOVERY

- Non-uniform schemes: $\forall \mathbf{x} \in \mathbb{R}^n$, $\Pr[\tilde{\mathbf{x}} \text{ recovers } \mathbf{x}] \geq 1 - \delta$
  - Count Sketch gives us such a guarantee
- Uniform schemes: $\Pr[\forall \mathbf{x} \in \mathbb{R}^n, \tilde{\mathbf{x}} \text{ recovers } \mathbf{x}] \geq 1 - \delta$

NON-UNIFORM VS. UNIFORM RECOVERY

- Non-uniform schemes: $\forall \mathbf{x} \in \mathbb{R}^n$, $\Pr[\tilde{\mathbf{x}} \text{ recovers } \mathbf{x}] \geq 1 - \delta$
  - Count Sketch gives us such a guarantee
- Uniform schemes: $\Pr[\forall \mathbf{x} \in \mathbb{R}^n, \tilde{\mathbf{x}} \text{ recovers } \mathbf{x}] \geq 1 - \delta$
  - Today: compressed sensing

NON-UNIFORM VS. UNIFORM RECOVERY

- Non-uniform schemes: $\forall \mathbf{x} \in \mathbb{R}^n$, $\Pr[\tilde{\mathbf{x}}$ recovers $\mathbf{x}] \geq 1 - \delta$
  - Count Sketch gives us such a guarantee
- Uniform schemes: $\Pr[\forall \mathbf{x} \in \mathbb{R}^n, \tilde{\mathbf{x}}$ recovers $\mathbf{x}] \geq 1 - \delta$
  - Today: compressed sensing
  - Pioneered by Candes & Tao

THE HALLMARK OF CS

THE HALLMARK OF CS

- A natural optimization problem: given $\mathbf{y} = A\mathbf{x}$, find $\tilde{\mathbf{x}}$, with $||\tilde{\mathbf{x}}||_0$ minimized, satisfying $A\tilde{\mathbf{x}} = \mathbf{y}$

THE HALLMARK OF CS

- A natural optimization problem: given $\mathbf{y} = A\mathbf{x}$, find $\tilde{\mathbf{x}}$, with $||\tilde{\mathbf{x}}||_0$ minimized, satisfying $A\tilde{\mathbf{x}} = \mathbf{y}$
  - This problem turns out NP-hard

THE HALLMARK OF CS

- A natural optimization problem: given $\mathbf{y} = A\mathbf{x}$, find $\tilde{\mathbf{x}}$, with $||\tilde{\mathbf{x}}||_0$ minimized, satisfying $A\tilde{\mathbf{x}} = \mathbf{y}$
  - This problem turns out NP-hard
- Compressed sensing solves the following linear program instead:

THE HALLMARK OF CS

- A natural optimization problem: given $\mathbf{y} = A\mathbf{x}$, find $\tilde{\mathbf{x}}$, with $||\tilde{\mathbf{x}}||_0$ minimized, satisfying $A\tilde{\mathbf{x}} = \mathbf{y}$
  - This problem turns out NP-hard
- Compressed sensing solves the following linear program instead:
  - $\min ||\hat{\mathbf{x}}||_1$, s.t., $A\hat{\mathbf{x}} = \mathbf{y}$. (*)

THE HALLMARK OF CS

- A natural optimization problem: given $\mathbf{y} = A\mathbf{x}$, find $\tilde{\mathbf{x}}$, with $||\tilde{\mathbf{x}}||_0$ minimized, satisfying $A\tilde{\mathbf{x}} = \mathbf{y}$
  - This problem turns out NP-hard
- Compressed sensing solves the following linear program instead:
  - $\min ||\hat{\mathbf{x}}||_1$, s.t., $A\hat{\mathbf{x}} = \mathbf{y}$. (*)
  - This is solvable in polynomial time

THE HALLMARK OF CS

- A natural optimization problem: given $\mathbf{y} = A\mathbf{x}$, find $\tilde{\mathbf{x}}$, with $||\tilde{\mathbf{x}}||_0$ minimized, satisfying $A\tilde{\mathbf{x}} = \mathbf{y}$
  - This problem turns out NP-hard
- Compressed sensing solves the following linear program instead:
  - $\min ||\hat{\mathbf{x}}||_1$, s.t., $A\hat{\mathbf{x}} = \mathbf{y}$. (*)
  - This is solvable in polynomial time
  - Intuitively, why is the solution to this LP a good recovery of $\mathbf{x}$?

RESTRICTED ISOMETRY PROPERTY

- In order for (*) to produce good recoveries, we need A to approximately preserve ℓ_2 norms for all sparse vectors.

Def. A matrix $A \in \mathbb{R}^{m \times n}$ is said to satisfy the *restricted isometry property* (RIP) with parameters α, β and s if the inequality

$$\alpha \|\mathbf{v}\|_2 \leq \|A\mathbf{v}\|_2 \leq \beta \|\mathbf{v}\|_2$$

holds for all vectors $\mathbf{v} \in \mathbb{R}^n$ such that $\|\mathbf{v}\|_0 \leq s$

RESTRICTED ISOMETRY PROPERTY

- In order for (*) to produce good recoveries, we need A to approximately preserve ℓ_2 norms for all sparse vectors.

Def. A matrix $A \in \mathbb{R}^{m \times n}$ is said to satisfy the *restricted isometry property* (RIP) with parameters α, β and s if the inequality

$$\alpha \|\mathbf{v}\|_2 \leq \|A\mathbf{v}\|_2 \leq \beta \|\mathbf{v}\|_2$$

holds for all vectors $\mathbf{v} \in \mathbb{R}^n$ such that $\|\mathbf{v}\|_0 \leq s$

Compare this with the distributional JL lemma

RANDOM MATRICES SATISFY RIP

RANDOM MATRICES SATISFY RIP

Theorem. Consider an $m \times n$ matrix A whose entries are i.i.d. drawn from the standard Gaussian $N(0,1)$. There are constants C and $c > 0$ such that, if $m \geq Cs\log(en/s)$, then with probability at least $1 - 2\exp(-cm)$, the random matrix A satisfies the RIP with parameters $\alpha = 0.9\sqrt{m}, \beta = 1.1\sqrt{m}$ and s.

RANDOM MATRICES SATISFY RIP

Theorem. Consider an $m \times n$ matrix A whose entries are i.i.d. drawn from the standard Gaussian $N(0,1)$. There are constants C and $c > 0$ such that, if $m \geq Cs\log(en/s)$, then with probability at least $1 - 2\exp(-cm)$, the random matrix A satisfies the RIP with parameters $\alpha = 0.9\sqrt{m}, \beta = 1.1\sqrt{m}$ and s.

RANDOM MATRICES SATISFY RIP

Theorem. Consider an $m \times n$ matrix A whose entries are i.i.d. drawn from the standard Gaussian $N(0,1)$. There are constants C and $c > 0$ such that, if $m \geq Cs\log(en/s)$, then with probability at least $1 - 2\exp(-cm)$, the random matrix A satisfies the RIP with parameters $\alpha = 0.9\sqrt{m}, \beta = 1.1\sqrt{m}$ and s.

The proof requires building up a bit of theory on random matrices.

RIP IMPLIES EXACT RECOVERY

RIP IMPLIES EXACT RECOVERY

Theorem. Suppose that an $m \times n$ matrix A satisfies the RIP with some parameters α, β and $(1+\lambda)s$, where $\lambda \geq (\beta/\alpha)^2$. Then every s-sparse vector $\mathbf{x} \in \mathbb{R}^n$ is recovered exactly by solving the program (*), i..e, the solution satisfies $\hat{\mathbf{x}} = \mathbf{x}$.

RIP IMPLIES EXACT RECOVERY

Theorem. Suppose that an $m \times n$ matrix A satisfies the RIP with some parameters α, β and $(1+\lambda)s$, where $\lambda \geq (\beta/\alpha)^2$. Then every s-sparse vector $\mathbf{x} \in \mathbb{R}^n$ is recovered exactly by solving the program (*), i..e, the solution satisfies $\hat{\mathbf{x}} = \mathbf{x}$.

Notations: Let $\mathbf{h} = \mathbf{x} - \hat{\mathbf{x}}$. We need to show $\mathbf{h} = \mathbf{0}$.

RIP IMPLIES EXACT RECOVERY

Theorem. Suppose that an $m \times n$ matrix A satisfies the RIP with some parameters α, β and $(1+\lambda)s$, where $\lambda \geq (\beta/\alpha)^2$. Then every s-sparse vector $\mathbf{x} \in \mathbb{R}^n$ is recovered exactly by solving the program (*), i..e, the solution satisfies $\hat{\mathbf{x}} = \mathbf{x}$.

Notations: Let $\mathbf{h} = \mathbf{x} - \hat{\mathbf{x}}$. We need to show $\mathbf{h} = \mathbf{0}$.

For any $T \subseteq [n]$, let $\mathbf{h}_T$ denote the vector $\mathbf{h}$ restricted to T, i.e., coordinates not in T are zeroed out

RIP IMPLIES EXACT RECOVERY

Theorem. Suppose that an $m \times n$ matrix A satisfies the RIP with some parameters α, β and $(1+\lambda)s$, where $\lambda \geq (\beta/\alpha)^2$. Then every s-sparse vector $\mathbf{x} \in \mathbb{R}^n$ is recovered exactly by solving the program (*), i..e, the solution satisfies $\hat{\mathbf{x}} = \mathbf{x}$.

Notations: Let $\mathbf{h} = \mathbf{x} - \hat{\mathbf{x}}$. We need to show $\mathbf{h} = \mathbf{0}$.

For any $T \subseteq [n]$, let $\mathbf{h}_T$ denote the vector $\mathbf{h}$ restricted to T, i.e., coordinates not in T are zeroed out

Let I_0 be the support of $\mathbf{x}$ (the set of non-zero coordinates), so $|I_0| \leq s$.

Lemma. $\|\mathbf{h}_{\overline{I_0}}\|_1 \leq \|\mathbf{h}_{I_0}\|_1$

Lemma. $\|\mathbf{h}_{\overline{I_0}}\|_1 \leq \|\mathbf{h}_{I_0}\|_1$

Proof. $\|\mathbf{x}\|_1 = \|\hat{\mathbf{x}}\|_1 = \|\mathbf{x} + \mathbf{h}\|_1 = \|\mathbf{x}_{I_0} + \mathbf{h}_{I_0}\|_1 + \|\mathbf{h}_{\overline{I_0}}\|_1 \geq \|\mathbf{x}\|_1 - \|\mathbf{h}_{I_0}\|_1 + \|\mathbf{h}_{\overline{I_0}}\|_1$

RIP IMPLIES EXACT RECOVERY

RIP IMPLIES EXACT RECOVERY

Theorem. Suppose that an $m \times n$ matrix A satisfies the RIP with some parameters α, β and $(1+\lambda)s$, where $\lambda \geq (\beta/\alpha)^2$. Then every s-sparse vector $\mathbf{x} \in \mathbb{R}^n$ is recovered exactly by solving the program (*), i..e, the solution satisfies $\hat{\mathbf{x}} = \mathbf{x}$.

RIP IMPLIES EXACT RECOVERY

Theorem. Suppose that an $m \times n$ matrix A satisfies the RIP with some parameters α, β and $(1+\lambda)s$, where $\lambda \geq (\beta/\alpha)^2$. Then every s-sparse vector $\mathbf{x} \in \mathbb{R}^n$ is recovered exactly by solving the program (*), i..e, the solution satisfies $\hat{\mathbf{x}} = \mathbf{x}$.

Notations: Let $\mathbf{h} = \mathbf{x} - \hat{\mathbf{x}}$. We need to show $\mathbf{h} = \mathbf{0}$.

RIP IMPLIES EXACT RECOVERY

Theorem. Suppose that an $m \times n$ matrix A satisfies the RIP with some parameters α, β and $(1+\lambda)s$, where $\lambda \geq (\beta/\alpha)^2$. Then every s-sparse vector $\mathbf{x} \in \mathbb{R}^n$ is recovered exactly by solving the program (*), i..e, the solution satisfies $\hat{\mathbf{x}} = \mathbf{x}$.

Notations: Let $\mathbf{h} = \mathbf{x} - \hat{\mathbf{x}}$. We need to show $\mathbf{h} = \mathbf{0}$.

For any $T \subseteq [n]$, let $\mathbf{h}_T$ denote the vector $\mathbf{h}$ restricted to T, i.e., coordinates not in T are zeroed out

RIP IMPLIES EXACT RECOVERY

Theorem. Suppose that an $m \times n$ matrix A satisfies the RIP with some parameters α, β and $(1+\lambda)s$, where $\lambda \geq (\beta/\alpha)^2$. Then every s-sparse vector $\mathbf{x} \in \mathbb{R}^n$ is recovered exactly by solving the program (*), i..e, the solution satisfies $\hat{\mathbf{x}} = \mathbf{x}$.

Notations: Let $\mathbf{h} = \mathbf{x} - \hat{\mathbf{x}}$. We need to show $\mathbf{h} = \mathbf{0}$.

For any $T \subseteq [n]$, let $\mathbf{h}_T$ denote the vector $\mathbf{h}$ restricted to T, i.e., coordinates not in T are zeroed out

Let I_0 be the support of $\mathbf{x}$ (the set of non-zero coordinates), so $|I_0| \leq s$.

RIP IMPLIES EXACT RECOVERY

Theorem. Suppose that an $m \times n$ matrix A satisfies the RIP with some parameters α, β and $(1+\lambda)s$, where $\lambda \geq (\beta/\alpha)^2$. Then every s-sparse vector $\mathbf{x} \in \mathbb{R}^n$ is recovered exactly by solving the program (*), i..e, the solution satisfies $\hat{\mathbf{x}} = \mathbf{x}$.

Notations: Let $\mathbf{h} = \mathbf{x} - \hat{\mathbf{x}}$. We need to show $\mathbf{h} = \mathbf{0}$.

For any $T \subseteq [n]$, let $\mathbf{h}_T$ denote the vector $\mathbf{h}$ restricted to T, i.e., coordinates not in T are zeroed out

Let I_0 be the support of $\mathbf{x}$ (the set of non-zero coordinates), so $|I_0| \leq s$.

Sort coordinates of $\mathbf{h}$ in $\overline{I_0}$ in absolute values. Let I_1 be the set of next largest λs coordinates, I_2 the next λs coordinates, and so on. Let $I_{0,1}$ be $I_0 \cup I_1$

Theorem. Suppose that an $m \times n$ matrix A satisfies the RIP with some parameters α, β and $(1+\lambda)s$, where $\lambda \geq (\beta/\alpha)^2$. Then every s-sparse vector $\mathbf{x} \in \mathbb{R}^n$ is recovered exactly by solving the program (*), i..e, the solution satisfies $\hat{\mathbf{x}} = \mathbf{x}$.

Theorem. Suppose that an $m \times n$ matrix A satisfies the RIP with some parameters α, β and $(1+\lambda)s$, where $\lambda \geq (\beta/\alpha)^2$. Then every s-sparse vector $\mathbf{x} \in \mathbb{R}^n$ is recovered exactly by solving the program (*), i..e, the solution satisfies $\hat{\mathbf{x}} = \mathbf{x}$.

Proof. We have $A\mathbf{h} = \mathbf{0}$, we have $0 = \|A\mathbf{h}\|_2 \geq \|A\mathbf{h}_{I_{0,1}}\|_2 - \|A\mathbf{h}_{\overline{I_{0,1}}}\|_2$

Theorem. Suppose that an $m \times n$ matrix A satisfies the RIP with some parameters α, β and $(1+\lambda)s$, where $\lambda \geq (\beta/\alpha)^2$. Then every s-sparse vector $\mathbf{x} \in \mathbb{R}^n$ is recovered exactly by solving the program (*), i..e, the solution satisfies $\hat{\mathbf{x}} = \mathbf{x}$.

Proof. We have $A\mathbf{h} = \mathbf{0}$, we have $0 = \|A\mathbf{h}\|_2 \geq \|A\mathbf{h}_{I_{0,1}}\|_2 - \|A\mathbf{h}_{\overline{I_{0,1}}}\|_2$

Since $|I_{0,1}| \leq s + \lambda s$, RIP yields $\|A\mathbf{h}_{I_{0,1}}\|_2 \geq \alpha \|\mathbf{h}_{I_{0,1}}\|_2$

Theorem. Suppose that an $m \times n$ matrix A satisfies the RIP with some parameters α, β and $(1+\lambda)s$, where $\lambda \geq (\beta/\alpha)^2$. Then every s-sparse vector $\mathbf{x} \in \mathbb{R}^n$ is recovered exactly by solving the program (*), i..e, the solution satisfies $\hat{\mathbf{x}} = \mathbf{x}$.

Proof. We have $A\mathbf{h} = \mathbf{0}$, we have $0 = \|A\mathbf{h}\|_2 \geq \|A\mathbf{h}_{I_{0,1}}\|_2 - \|A\mathbf{h}_{\overline{I_{0,1}}}\|_2$

Since $|I_{0,1}| \leq s + \lambda s$, RIP yields $\|A\mathbf{h}_{I_{0,1}}\|_2 \geq \alpha \|\mathbf{h}_{I_{0,1}}\|_2$

On the other side, $\|A\mathbf{h}_{\overline{I_{0,1}}}\|_2 \leq \sum_{i\geq 2} \|A\mathbf{h}_{I_i}\|_2 \leq \beta \sum_{i \geq 2} \|\mathbf{h}_{I_i}\|_2$

Theorem. Suppose that an $m \times n$ matrix A satisfies the RIP with some parameters α, β and $(1+\lambda)s$, where $\lambda \geq (\beta/\alpha)^2$. Then every s-sparse vector $\mathbf{x} \in \mathbb{R}^n$ is recovered exactly by solving the program (*), i..e, the solution satisfies $\hat{\mathbf{x}} = \mathbf{x}$.

Proof. We have $A\mathbf{h} = \mathbf{0}$, we have $0 = \|A\mathbf{h}\|_2 \geq \|A\mathbf{h}_{I_{0,1}}\|_2 - \|A\mathbf{h}_{\overline{I_{0,1}}}\|_2$

Since $|I_{0,1}| \leq s + \lambda s$, RIP yields $\|A\mathbf{h}_{I_{0,1}}\|_2 \geq \alpha \|\mathbf{h}_{I_{0,1}}\|_2$

On the other side, $\|A\mathbf{h}_{\overline{I_{0,1}}}\|_2 \leq \sum_{i \geq 2} \|A\mathbf{h}_{I_i}\|_2 \leq \beta \sum_{i \geq 2} \|\mathbf{h}_{I_i}\|_2$

Crucial step: for $i \geq 2$, $\forall j \in I_i$, we have $|h_j| \leq \frac{1}{\lambda s} \|\mathbf{h}_{I_{i-1}}\|_1$, and so $\|\mathbf{h}_{I_i}\|_2 \leq \frac{1}{\sqrt{\lambda s}} \|\mathbf{h}_{I_{i-1}}\|_1$

Theorem. Suppose that an $m \times n$ matrix A satisfies the RIP with some parameters α, β and $(1+\lambda)s$, where $\lambda \geq (\beta/\alpha)^2$. Then every s-sparse vector $\mathbf{x} \in \mathbb{R}^n$ is recovered exactly by solving the program (*), i..e, the solution satisfies $\hat{\mathbf{x}} = \mathbf{x}$.

Proof. We have $A\mathbf{h} = \mathbf{0}$, we have $0 = \|A\mathbf{h}\|_2 \geq \|A\mathbf{h}_{I_{0,1}}\|_2 - \|A\mathbf{h}_{\overline{I_{0,1}}}\|_2$

Since $|I_{0,1}| \leq s + \lambda s$, RIP yields $\|A\mathbf{h}_{I_{0,1}}\|_2 \geq \alpha \|\mathbf{h}_{I_{0,1}}\|_2$

On the other side, $\|A\mathbf{h}_{\overline{I_{0,1}}}\|_2 \leq \sum_{i\geq 2} \|A\mathbf{h}_{I_i}\|_2 \leq \beta \sum_{i\geq 2} \|\mathbf{h}_{I_i}\|_2$

Crucial step: for $i \geq 2$, $\forall j \in I_i$, we have $|h_j| \leq \frac{1}{\lambda s}\|\mathbf{h}_{I_{i-1}}\|_1$, and so $\|\mathbf{h}_{I_i}\|_2 \leq \frac{1}{\sqrt{\lambda s}}\|\mathbf{h}_{I_{i-1}}\|_1$

Therefore $\sum_{i\geq 2} \|\mathbf{h}_{I_i}\|_2 \leq \frac{1}{\sqrt{\lambda s}} \sum_{i\geq 1} \|\mathbf{h}_{I_i}\|_1 = \frac{1}{\sqrt{\lambda s}}\|\mathbf{h}_{\overline{I_0}}\|_1 \leq \frac{1}{\sqrt{\lambda s}}\|\mathbf{h}_{I_0}\|_1 \leq \frac{1}{\sqrt{\lambda}}\|\mathbf{h}_{I_0}\|_2 \leq \frac{1}{\sqrt{\lambda}}\|\mathbf{h}_{I_{0,1}}\|_2$

Theorem. Suppose that an $m \times n$ matrix A satisfies the RIP with some parameters α, β and $(1+\lambda)s$, where $\lambda \geq (\beta/\alpha)^2$. Then every s-sparse vector $\mathbf{x} \in \mathbb{R}^n$ is recovered exactly by solving the program (*), i..e, the solution satisfies $\hat{\mathbf{x}} = \mathbf{x}$.

Proof. We have $A\mathbf{h} = \mathbf{0}$, we have $0 = \|A\mathbf{h}\|_2 \geq \|A\mathbf{h}_{I_{0,1}}\|_2 - \|A\mathbf{h}_{\overline{I_{0,1}}}\|_2$

Since $|I_{0,1}| \leq s + \lambda s$, RIP yields $\|A\mathbf{h}_{I_{0,1}}\|_2 \geq \alpha \|\mathbf{h}_{I_{0,1}}\|_2$

On the other side, $\|A\mathbf{h}_{\overline{I_{0,1}}}\|_2 \leq \sum_{i \geq 2} \|A\mathbf{h}_{I_i}\|_2 \leq \beta \sum_{i \geq 2} \|\mathbf{h}_{I_i}\|_2$

Theorem. Suppose that an $m \times n$ matrix A satisfies the RIP with some parameters α, β and $(1+\lambda)s$, where $\lambda \geq (\beta/\alpha)^2$. Then every s-sparse vector $\mathbf{x} \in \mathbb{R}^n$ is recovered exactly by solving the program (*), i..e, the solution satisfies $\hat{\mathbf{x}} = \mathbf{x}$.

Proof. We have $A\mathbf{h} = \mathbf{0}$, we have $0 = \|A\mathbf{h}\|_2 \geq \|A\mathbf{h}_{I_{0,1}}\|_2 - \|A\mathbf{h}_{\overline{I_{0,1}}}\|_2$

Since $|I_{0,1}| \leq s + \lambda s$, RIP yields $\|A\mathbf{h}_{I_{0,1}}\|_2 \geq \alpha \|\mathbf{h}_{I_{0,1}}\|_2$

On the other side, $\|A\mathbf{h}_{\overline{I_{0,1}}}\|_2 \leq \sum_{i \geq 2} \|A\mathbf{h}_{I_i}\|_2 \leq \beta \sum_{i \geq 2} \|\mathbf{h}_{I_i}\|_2$

So we have $\frac{\beta}{\sqrt{\lambda}} \|\mathbf{h}_{I_{0,1}}\|_2 \geq \alpha \|\mathbf{h}_{I_{0,1}}\|_2$. But since $\lambda \geq (\beta/\alpha)^2$, this implies $\mathbf{h}_{I_{0,1}} = \mathbf{0}$, which by definition means $\mathbf{h} = \mathbf{0}$.