Sparse Recovery

Dhanus M Lal

April 26, 2022

Introduction

- The number of linear measurements required to completely understand a signal in $\mathbb{R}^n$ is n.

Introduction

- The number of linear measurements required to completely understand a signal in $\mathbb{R}^n$ is n.
- If it is known that all but the first coordinate of it is zero, only one measurement, $y = \langle e_1, x \rangle$ is required.

Introduction

- The number of linear measurements required to completely understand a signal in $\mathbb{R}^n$ is n.
- If it is known that all but the first coordinate of it is zero, only one measurement, $y = \langle e_1, x \rangle$ is required.
- If we know that at most one coordinate is non-zero, can we learn the signal with fewer than n measurements?

Introduction

- The number of linear measurements required to completely understand a signal in $\mathbb{R}^n$ is n.
- If it is known that all but the first coordinate of it is zero, only one measurement, $y = \langle e_1, x \rangle$ is required.
- If we know that at most one coordinate is non-zero, can we learn the signal with fewer than n measurements?
- Surprising result is that such a signal can be recovered in $\mathcal{O}(\log n)$ measurements

Introduction

- The number of linear measurements required to completely understand a signal in $\mathbb{R}^n$ is n.
- If it is known that all but the first coordinate of it is zero, only one measurement, $y = \langle e_1, x \rangle$ is required.
- If we know that at most one coordinate is non-zero, can we learn the signal with fewer than n measurements?
- Surprising result is that such a signal can be recovered in $\mathcal{O}(\log n)$ measurements

Definition (Sparsity)

The number of non-zero coordinates of a signal in $\mathbb{R}^n$. Denoted by $\|x\|_0$.

Sparse Recovery Problem

- Recover x from linear measurement $y = Ax$ with x being $s-$sparse and $A \in \mathbb{R}^{m \times n}$ known.

Sparse Recovery Problem

- Recover x from linear measurement $y = Ax$ with x being $s-$sparse and $A \in \mathbb{R}^{m \times n}$ known.
- Is there a measurement matrix $A \in \mathbb{R}^{m \times n}$ that makes recovery possible?

Sparse Recovery Problem

- Recover x from linear measurement $y = Ax$ with x being $s-$sparse and $A \in \mathbb{R}^{m \times n}$ known.
- Is there a measurement matrix $A \in \mathbb{R}^{m \times n}$ that makes recovery possible?
- Answer is yes!

Sparse Recovery Problem

- Recover x from linear measurement $y = Ax$ with x being s – sparse and $A \in \mathbb{R}^{m \times n}$ known.
- Is there a measurement matrix $A \in \mathbb{R}^{m \times n}$ that makes recovery possible?
- Answer is yes!

Proposition

If $A \in \mathbb{R}^{m \times n}$ has $m > 2s$ and every subset of $2s$ columns of A is linearly independent, then sparse recovery problem has a unique solution.

Sparse Recovery Problem

- Recover x from linear measurement $y = Ax$ with x being $s-$ sparse and $A \in \mathbb{R}^{m \times n}$ known.
- Is there a measurement matrix $A \in \mathbb{R}^{m \times n}$ that makes recovery possible?
- Answer is yes!

Proposition

If $A \in \mathbb{R}^{m \times n}$ has $m > 2s$ and every subset of $2s$ columns of A is linearly independent, then sparse recovery problem has a unique solution.

Proof.

If x_1 and x_2 are two s–sparse vectors with $Ax_1 = Ax_2$, then $A(x_1 - x_2) = 0$. Now, $x_1 \neq x_2$ would mean that a subset of $2s$ columns is linearly dependent. $\square$

Basis Pursuit

- If solution is unique, the following optimization problem gives the actual signal.

$$x = \arg\min \|x'\|_0 \quad \text{s.t } Ax' = y$$

Basis Pursuit

- If solution is unique, the following optimization problem gives the actual signal.

$$x = \arg\min \|x'\|_0 \quad \text{s.t } Ax' = y$$

- Solving this is computationally hard: requires solving $\binom{n}{s}$ linear systems.

Basis Pursuit

- If solution is unique, the following optimization problem gives the actual signal.

$$x = \arg\min \|x'\|_0 \quad \text{s.t } Ax' = y$$

- Solving this is computationally hard: requires solving $\binom{n}{s}$ linear systems.
- Solve the convex relaxation instead

$$\hat{x} = \arg\min \|x'\|_1 \quad \text{s.t } Ax' = y$$

Basis Pursuit

- If solution is unique, the following optimization problem gives the actual signal.

$$x = \arg\min \|x'\|_0 \quad \text{s.t } Ax' = y$$

- Solving this is computationally hard: requires solving $\binom{n}{s}$ linear systems.
- Solve the convex relaxation instead

$$\hat{x} = \arg\min \|x'\|_1 \quad \text{s.t } Ax' = y$$

- This is a convex optimization problem called basis pursuit, it is computationally efficient.

Basis Pursuit

- If solution is unique, the following optimization problem gives the actual signal.

$$x = \arg\min \|x'\|_0 \quad \text{s.t } Ax' = y$$

- Solving this is computationally hard: requires solving $\binom{n}{s}$ linear systems.
- Solve the convex relaxation instead

$$\hat{x} = \arg\min \|x'\|_1 \quad \text{s.t } Ax' = y$$

- This is a convex optimization problem called basis pursuit, it is computationally efficient.
- Questions:
    - How close is $\hat{x}$ to x?
    - Exact recovery when?

Recovery Using Random Matrices

Definition

$\mathcal{A}(m, n)$ is the class of random matrices with isotropic, sub-gaussian, independent rows.

Recovery Using Random Matrices

Definition

$\mathcal{A}(m, n)$ is the class of random matrices with isotropic, sub-gaussian, independent rows.

Several useful properties of $\mathcal{A}(m, n)$ are known.

Recovery Using Random Matrices

Definition

$\mathcal{A}(m, n)$ is the class of random matrices with isotropic, sub-gaussian, independent rows.

Several useful properties of $\mathcal{A}(m, n)$ are known.

Theorem

If $A \in \mathcal{A}(m, n)$ and $E = A^{-1}(z)$, then for any subset T of $\mathbb{R}^n$,

$$\mathbb{E} diam(T \cap E) \leq Cw(T)/\sqrt{m}$$

Recovery Using Random Matrices

Definition

$\mathcal{A}(m, n)$ is the class of random matrices with isotropic, sub-gaussian, independent rows.

Several useful properties of $\mathcal{A}(m, n)$ are known.

Theorem

If $A \in \mathcal{A}(m, n)$ and $E = A^{-1}(z)$, then for any subset T of $\mathbb{R}^n$,

$$\mathbb{E} diam(T \cap E) \leq Cw(T)/\sqrt{m}$$

This can be used to get

$$\mathbb{E}\|x - \hat{x}\|_2 \leq C\sqrt{\frac{s \log n}{m}}\|x\|_2$$

Exact Recovery for $A \in \mathcal{A}$

Another property of matrices in $\mathcal{A}$.

Theorem

If $A \in \mathcal{A}(m, n)$ with $m \geq Cw(T)^2$, for $T \subset S^{n-1}$, then $T \cap \ker(A) = \emptyset$ with high probability ($\geq 1 - 2\exp(-cm)$)

Exact Recovery for $A \in \mathcal{A}$

Another property of matrices in $\mathcal{A}$.

Theorem

If $A \in \mathcal{A}(m, n)$ with $m \geq Cw(T)^2$, for $T \subset S^{n-1}$, then $T \cap \ker(A) = \emptyset$ with high probability ($\geq 1 - 2\exp(-cm)$)

- From theorem, $\hat{x} = x$ with high probability provided $m > Cs \log n$.
- Because, if $h = \hat{x} - x \neq 0$, then $h/\|h\|_2 \in S^{n-1} \cap 2\sqrt{s}B_1^n$, which is empty with high probability.

Exact Recovery for $A \in \mathcal{A}$

Another property of matrices in $\mathcal{A}$.

Theorem

If $A \in \mathcal{A}(m, n)$ with $m \geq Cw(T)^2$, for $T \subset S^{n-1}$, then $T \cap \ker(A) = \emptyset$ with high probability ($\geq 1 - 2\exp(-cm)$)

- From theorem, $\hat{x} = x$ with high probability provided $m > Cs \log n$.
- Because, if $h = \hat{x} - x \neq 0$, then $h/\|h\|_2 \in S^{n-1} \cap 2\sqrt{s}B_1^n$, which is empty with high probability.
- Exact recovery using basis pursuit is possible!

RIP Implies Exact Recovery

Definition (RIP)

A matrix A satisfies RIP with parameter s if the following holds for every $3s$–sparse vectors:

$$0.9\|x\|_2 \leq \|Ax\|_2 \leq 1.1\|x\|_2$$

RIP Implies Exact Recovery

Definition (RIP)

A matrix A satisfies RIP with parameter s if the following holds for every $3s$–sparse vectors:

$$0.9\|x\|_2 \leq \|Ax\|_2 \leq 1.1\|x\|_2$$

- RIP implies exact recovery for s–sparse vectors. Proof of this only involves triangle inequality and Cauchy-Schwarz inequality.

RIP Implies Exact Recovery

Definition (RIP)

A matrix A satisfies RIP with parameter s if the following holds for every $3s$–sparse vectors:

$$0.9\|x\|_2 \leq \|Ax\|_2 \leq 1.1\|x\|_2$$

- RIP implies exact recovery for s–sparse vectors. Proof of this only involves triangle inequality and Cauchy-Schwarz inequality.
- All $m \times 3s$ sub-matrices must be almost isometry: difficult condition.

RIP Implies Exact Recovery

Definition (RIP)

A matrix A satisfies RIP with parameter s if the following holds for every $3s$–sparse vectors:

$$0.9\|x\|_2 \leq \|Ax\|_2 \leq 1.1\|x\|_2$$

- RIP implies exact recovery for s–sparse vectors. Proof of this only involves triangle inequality and Cauchy-Schwarz inequality.
- All $m \times 3s$ sub-matrices must be almost isometry: difficult condition.
- Surprisingly, $A \in \mathcal{A}(m, n)$ satisfies RIP if $m > Cs \log n$ with high probability ($\geq 1 - 2\exp(-cm)$).
- Alternate proof of the previous result.

Further Reading

- Recovery of sparse signal using partial Fourier coefficients.

Further Reading

- Recovery of sparse signal using partial Fourier coefficients.
- Low rank matrix recovery

Further Reading

- Recovery of sparse signal using partial Fourier coefficients.
- Low rank matrix recovery
- Numerical simulations.