Approximation Algorithms for Model-Based Compressive Sensing

Jaron Chen

Agenda

- Motivation
- Introduction to Approximate Model Algorithms
- Mathematical Background
- Approximate Model Iterative Hard Thresholding
- Approximate Model Compressive Sampling Matching Pursuit
- Improved Recovery via Boosting
- Summary

Motivation

Compressive sensing allows for a sparse signal to be recovered from a limited number of measurements

Model-based compressive sensing uses the structure of the signal to recover the signal with even fewer measurements

Motivation: Robust Sparse Recovery

A is an m-by-n matrix.

x is the original n-dimensional k-sparse signal.

e is the noise vector.

$$y = Ax + e$$
$$\|x - \hat{x}\|_2 \leq C\|e\|_2$$

where C is the constant approximation factor.

The number of measurements needed: $\mathrm{m} = O\left(k \log \frac{\mathrm{n}}{k}\right)$.

Large n can cause m to be very large.

Introduction to Approximate Model Algorithms

Approximation-Tolerant Model-Based Compressive Sensing.

Using careful signal modeling can overcome this limitation. One way is through a method known as approximation-tolerant model-based compressive sensing. This framework includes sparse-recovery algorithms that only require approximate solutions.

This can reduce the number of measurements needed to recover the signal.

Mathematical Background

Mathematical Definitions

Let [n] denote the set {1,2,…,n} and $\Omega \subseteq [n]$.

$$(x_\Omega)_i = \begin{cases} x_i, & i \in \Omega \\ 0, & \text{otherwise} \end{cases}$$

For a matrix $A \in \mathbb{R}^{\mathrm{m}\times\mathrm{n}}$,

A_Ω is the submatrix with the columns corresponding to Ω ($A_\Omega \in \mathbb{R}^{\mathrm{m}\times|\Omega|}$).

Structured Sparsity Model

Let the structured sparsity model $M \subseteq \mathbb{R}^{\mathrm{n}}$, which is the set of vectors $\boldsymbol{\mathcal{M}}= \{x \in \mathbb{R}^{\mathrm{n}} \mid \mathrm{supp}(x) \subseteq S$ for some $S \in \mathbb{M}\}$, where $\mathbb{M}$=$\{\Omega_1, \dots, \Omega_l\}$ and l is the size of $\boldsymbol{\mathcal{M}}$.

Let $\mathbb{M}^{+} = \{\Omega \subseteq [\mathrm{n}] \mid \Omega \subseteq S \text{ for some } S \in \mathbb{M}\}$

Therefore, the set of vectors $\boldsymbol{\mathcal{M}}= \{x \in \mathbb{R}^{\mathrm{n}} \mid \mathrm{supp}(x) \in \mathbb{M}^{+}\}$

Model Restricted Isometry Property

The matrix $A \in \mathbb{R}^{\mathrm{m}\times\mathrm{n}}$ satisfies the $(\delta, \mathbb{M})$-model-RIP if this equality holds for all x with $\mathrm{supp}(x) \in \mathbb{M}^{+}$.

$$(1-\delta)\|x\|_2^2 \leq \|Ax\|_2^2 \leq (1+\delta)\|x\|_2^2$$

Let $A \in \mathbb{R}^{\mathrm{m}\times\mathrm{n}}$ satisfy the RIP. Let $\Omega \in \mathbb{M}^{+}, x \in \mathbb{R}^{\mathrm{n}}$, and $\mathrm{y} \in \mathbb{R}^{\mathrm{m}}$. The following properties will hold:

$$\left\|A_\Omega^T y\right\|_2 \leq \sqrt{1+\delta}\|y\|_2$$

$$\left\|A_\Omega^T A_\Omega x\right\|_2 \leq \sqrt{1+\delta}\|x\|_2$$

$$\left\|(I - A_\Omega^T A_\Omega)x\right\|_2 \leq \sqrt{1+\delta}\|x\|_2$$

Model-projection oracle

Model-projection oracle is a function $M\colon \mathbb{R}^{\mathrm{n}} \longrightarrow \mathcal{P}\left([\mathrm{n}]\right)$ that follows the output model sparsity and optimal model projection properties.

Output model sparsity: $M(x) \in \mathbb{M}^{+}$.

Optimal model projection:

Let $\Omega' = M(x)$. Then, $\|x - x_{\Omega'}\|_2 = \min_{\Omega \in \mathbb{M}} \|x - x_{\Omega}\|_2$.

Model Iterative Hard Thresholding

One of the most popular algorithms for sparse recovery is iterative hard thresholding (IHT). This can be modified to apply to the model $\boldsymbol{\mathcal{M}}$.

$$x^{i+1} \leftarrow M\left(x^i + A^T\left(y - Ax^i\right)\right)$$

where x^1, the initial signal estimate, is 0. This is executed until convergence.

Measurement Bound

Let k=$\max_{\Omega\in\mathbb{M}}|\Omega|$ and $0 < \delta < 1$.

There must be a constant c such that m$\geq \frac{c}{\delta^2}\left(k \log\frac{1}{\delta} + \log|\mathbb{M}| + t\right)$ for any $t > 0$.

m$= O(k + \log|\mathbb{M}|) = O(k)$

Incorrect Approach

For structured sparsity models, computing the optimal model-projection can be difficult. This can be simplified by using approximate model-projection oracles.

In a standard compressive sensing setting, the model consists of the set of all k-sparse signals. Thus, the oracle $\mathrm{T}_k(\cdot)$ returns the k coefficients with the largest magnitude of x.

Let c be an arbitrary constant and T_k' be a projection oracle such that for any $a \in \mathbb{R}$:

$$\|a - T_k'(a)\|_2 \leq c\|a - T_k(a)\|$$

Incorrect Approach

Adapting this to the Model-IHT,

$$x^{i+1} \leftarrow T'_k\left(x^i + A^T\left(y - Ax^i\right)\right)$$

where the first iteration is $x^1 \leftarrow T'_k(A^T y)$

Why is this approach incorrect?

Consider a 1-sparse signal x with $x_1 = 1$ and $x_i = 0$ for $i \neq 1$ and a matrix A with $\left(\delta, O(1)\right)$-RIP for small δ.

$$x^{i+1} \leftarrow T_k'\left(x^i + A^T\left(y - Ax^i\right)\right)$$

$$x^1 = x^0 = 0$$

Algorithms Assumptions

The algorithms use two projection oracles. Given $x \in \mathbb{R}^{\mathrm{n}}$, a tail approximation oracle returns a support Ω_t in the model such that the norm of the tail $\left\|x - x_{\Omega_t}\right\|_2$ is approximately minimized. A head approximation oracle returns a support Ω_h in the model such that the norm of the tail $\left\|x_{\Omega_t}\right\|_2$ is approximately minimized.

Approximate Oracles

Head Approximation Oracle

Let $\mathbb{M}, \mathbb{M}_H \subseteq \mathcal{P}([\mathrm{n}]), p \geq 1, \text{and } c_H \in \mathbb{R}$.

$H\colon\ \mathbb{R}^{\mathrm{n}} \longrightarrow \mathcal{P}([\mathrm{n}])$ is a $(c_H, \mathbb{M}, \mathbb{M}_H, p)$-head approximation oracle if output model sparsity and optimal model projection properties hold.

Output model sparsity: $H(x) \in \mathbb{M}_H^+$.

Head approximation:

$\text{Let } \Omega' = H(x)$. Then, $\|x_{\Omega'}\|_p \geq\ c_H \|x_\Omega\|_p \text{ for all } \Omega \in \mathbb{M}$.

Approximation Oracles

Tail Approximation Oracle

Let $\mathbb{M}, \mathbb{M}_T \subseteq \mathcal{P}([\mathrm{n}]), p \geq 1, \text{and } c_T \in \mathbb{R}$.

T: $\mathbb{R}^{\mathrm{n}} \longrightarrow \mathcal{P}([\mathrm{n}])$ is a $(c_T, \mathbb{M}, \mathbb{M}_T, p)$-tail approximation oracle if output model sparsity and optimal model projection properties hold.

Output model sparsity: $T(x) \in \mathbb{M}_T^+$.

Tail approximation:

Let $\Omega' = T(x)$. Then, $\|x - x_{\Omega'}\|_p \geq\ c_T\|x - x_\Omega\|_p$ for all $\Omega \in \mathbb{M}$.

Approximate Algorithms

Approximate Model Iterative Hard Thresholding

Approximate Model-IHT Algorithm

Algorithm 1 Approximate Model-IHT

function AM-IHT(y, A, t)
    $x^0 \leftarrow 0$
    for $i \leftarrow 0, \ldots, t$ **do**
        $b^i \leftarrow A^T(y - Ax^i)$
        $x^{i+1} \leftarrow T(x^i + H(b^i))$
    return x^{t+1}

Assumptions on Algorithm

1. $x \in \mathbb{R}^{\mathrm{n}}$ and $x \in \boldsymbol{\mathcal{M}}$

2. $y = Ax + e$ for $e \in \mathbb{R}^{\mathrm{m}}$

3. T is a $(c_T, \mathbb{M}, \mathbb{M}_T, 2)$-tail approximation oracle.

4. H is a $(c_H, \mathbb{M}_T \oplus \mathbb{M}, \mathbb{M}_H, 2)$-head approximation oracle.

5. A has the $(\delta,\, \mathbb{M} \bigoplus \mathbb{M}_T \bigoplus \mathbb{M}_H)$-model RIP.

Let the sum $\mathbb{C} = \mathbb{A} \bigoplus \mathbb{B} = \{\Omega + \Gamma \,|\Omega \in \mathbb{A} \text{ and } \Gamma \in \mathbb{B}\}$

Model-RIP on relevant vectors

Let $r^i = x - x^i$, $\Omega = \mathrm{supp}\left(r^i\right)$, and $\Gamma = \mathrm{supp}(H\left(b^i\right))$. For all $x' \in \mathbb{R}^{\mathrm{n}}$ with $\mathrm{supp}(x') \subseteq \Omega \cup T$,

$$(1-\delta)\|x'\|_2^2 \leq \|Ax'\|_2^2 \leq (1+\delta)\|x\|_2^2$$

Proof: Because $\mathrm{supp}\left(x^i\right) \in \mathbb{M}_T$ and $\mathrm{supp}(x) \in \mathbb{M}$, $\mathrm{supp}\left(x - x^i\right) \in \mathbb{M}_T \bigoplus \mathbb{M}$, and therefore, $\Omega \in \mathbb{M}_T \bigoplus \mathbb{M}$. $\mathrm{supp}\left(H\left(b^i\right)\right) \in \mathbb{M}_H$, so $\Omega \cup \Gamma \in \mathbb{M} \bigoplus \mathbb{M}_T \bigoplus \mathbb{M}_H$, which allows model-RIP to be performed.

Geometric Convergence

Let $r^i = x - x^i$ where x^i is the signal estimate at iteration i.

$$\left\|r^{i+1}\right\|_2 \leq \alpha\left\|r^i\right\|_2 + \beta\|e\|_2$$

where $\alpha = (1 + c_T)\left(\delta + \sqrt{1 - \alpha_0^2}\right)$,

$$\beta = (1 + c_T)\left(\frac{\beta_0}{\alpha_0} + \frac{\alpha_0 \beta_0}{\sqrt{1-\alpha_0^2}} + \sqrt{1 + \delta}\right),$$

$\alpha_0 = c_H(1 - \delta) - \delta$ and $\beta_0 = (1 + c_H)\sqrt{1 + \delta}$

Proof of Geometric Convergence

$$a = x^i + H\left(b^i\right)$$

Using the triangle inequality,

$$\begin{aligned}
\left\|x - x^{i+1}\right\|_2 &= \|x - T(a)\|_2 \\
&\leq \|x - a\|_2 + \|a - T(a)\|_2 \\
&\leq (1 + c_T)\|x - a\|_2 \\
&= (1 + c_T)\left\|x - x^i - H(b^i)\right\|_2 \\
&= (1 + c_T)\left\|r^i - H(A^T A r^i + A^T e)\right\|_2
\end{aligned}$$

Lemma Proof

Let $\Omega = \operatorname{supp}\left(r^i\right)$ and $\Gamma = \operatorname{supp}(H\left(b^i\right))$.

$$\left\|r_{\Gamma^c}^i\right\| \leq \sqrt{1-\alpha_0^2}\left\|r^i\right\|_2 + \left[\frac{\beta_0}{\alpha_0} + \frac{\alpha_0\beta_0}{\sqrt{1-\alpha_0^2}}\right]\|e\|_2$$

where $\alpha_0 = c_H(1-\delta) - \delta$ and $\beta_0 = (1+c_H)\sqrt{1+\delta}$

Lemma Proof: Lower Bound on $\|b^i_\Gamma\|_2$

$$b^i = A^T(y - Ax_i) = A^T A x^i + A^T e$$

Using the head-approximation and triangle inequality properties,

$$\begin{aligned}
\|b^i_\Gamma\|_2 &= \|A^T_\Gamma A r^i + A^T_\Gamma e\|_2 \\
&\geq c_H \|A^T_\Omega A r^i + A^T_\Omega e\|_2 \\
&\geq c_H \|A^T_\Omega A_\Omega r^i\|_2 - c_H \|A^T_\Omega e\|_2 \\
&\geq c_H (1 - \delta)\|r^i\|_2 - c_H \sqrt{1 + \delta}\|e\|_2
\end{aligned}$$

Lemma Proof: Upper Bound on $\|b_\Gamma^i\|_2$

$$b^i = A^T(y - Ax_i) = A^T A x^i + A^T e$$

Using the triangle inequality property and restricted isometry property,

$$\begin{aligned}
\|b_\Gamma^i\|_2 &= \|A_\Gamma^T A r^i + A_\Gamma^T e\|_2 \\
&= \|A_\Gamma^T A r^i - r_\Gamma^i + r_\Gamma^i + A_\Gamma^T e\|_2 \\
&\leq \|A_\Gamma^T A r^i - r_\Gamma^i\|_2 + \|r_\Gamma^i\|_2 + \|A_\Gamma^T e\|_2 \\
&\leq \|A_{\Gamma\cup\Omega}^T A r^i - r_{\Gamma\cup\Omega}^i\|_2 + \|r_\Gamma^i\|_2 + \sqrt{1+\delta}\|e\|_2 \\
&\leq \delta\|r^i\|_2 + \|r_\Gamma^i\|_2 + \sqrt{1+\delta}\|e\|_2
\end{aligned}$$

Lemma Proof

Combining the lower and upper bounds on $\left\|b_\Gamma^i\right\|_2$,

$$\left\|r_\Gamma^i\right\| \geq \alpha_0 \left\|r^i\right\|_2 - \beta_0 \|e\|_2$$

where $\alpha_0 = c_H(1-\delta) - \delta$ and $\beta_0 = (1 + c_H)\sqrt{1+\delta}$

Lemma Proof: Case 1

If $\alpha_0 \left\| r^i \right\|_2 \leq \beta_0 \|e\|_2$,

$$\left\| r_{\Gamma^c}^i \right\|_2 \leq \frac{\beta_0}{\alpha_0} \|e\|_2$$

because $\left\| r^i \right\|_2 > \left\| r_{\Gamma^c}^i \right\|_2$.

Lemma Proof: Case 2

If $\alpha_0 \left\| r^i \right\|_2 \geq \beta_0 \|e\|_2$,

$$\left\| r_\Gamma^i \right\|_2 \geq \left\| r^i \right\|_2 \left(\alpha_0 - \beta_0 \frac{\|e\|_2}{\|r^i\|_2} \right)$$

Knowing $\left\| r^i \right\|_2 = \left\| r_\Gamma^i \right\|_2 + \left\| r_{\Gamma^c}^i \right\|_2$,

$$\left\| r_{\Gamma^c}^i \right\|_2 \leq \left\| r^i \right\|_2 \sqrt{1 - \left(\alpha_0 - \beta_0 \frac{\|e\|_2}{\|r^i\|_2} \right)^2}$$

Lemma Proof: Case 2 (continued)

The $\sqrt{1-\left(\alpha_0-\beta_0\frac{\|e\|_2}{\|r^i\|_2}\right)^2}$ term can be reduced.

$$\omega_0=\alpha_0-\beta_0\frac{\|e\|_2}{\|r^i\|_2}$$

$\omega_0<1$ because $\alpha_0\|r^i\|_2\geq\beta_0\|e\|_2, \alpha_0<1,$ and $\beta_0\geq 1.$

If $0<\omega<1$,

$$\sqrt{1-\omega_0^2}\leq\frac{1}{\sqrt{1-\omega^2}}-\frac{\omega}{\sqrt{1-\omega^2}}\omega_0$$

Lemma Proof: Case 2 (continued)

$$\left\|r_{\Gamma^c}^i\right\|_2 \leq \left\|r^i\right\|_2 \left(\frac{1}{\sqrt{1-\omega^2}} - \frac{\omega}{\sqrt{1-\omega^2}}\left(\alpha_0 - \beta_0 \frac{\|e\|_2}{\|r^i\|_2}\right)\right)$$

$$\frac{1-\omega\alpha_0}{\sqrt{1-\omega^2}}\left\|r^i\right\|_2 + \frac{\omega\beta_0}{\sqrt{1-\omega^2}}\|e\|_2$$

$\frac{1-\omega\alpha_0}{\sqrt{1-\omega^2}}$ determines the convergence rate, and this is minimized if $\omega = \alpha_0$.

$$\left\|r_{\Gamma^c}^i\right\|_2 \leq \sqrt{1-\alpha_0^2}\left\|r^i\right\|_2 + \frac{\alpha_0\beta_0}{\sqrt{1-\alpha_0^2}}\|e\|_2$$

Lemma Proof (continued)

Combining the results from cases 1 and 2,

$$\left\|r_{\Gamma^c}^i\right\| \le \sqrt{1-\alpha_0^2}\left\|r^i\right\|_2 + \left[\frac{\beta_0}{\alpha_0} + \frac{\alpha_0\beta_0}{\sqrt{1-\alpha_0^2}}\right]\|e\|_2$$

Proof of Geometric Convergence (continued)

$\left\| r^i - H(A^T A r^i + A^T e) \right\|_2$ can be bounded.

Let $\Omega = \operatorname{supp}(r^i)$ and $\Gamma = \operatorname{supp}\left(H\left(A^T A r^i + A^T e\right)\right)$.

$$
\begin{aligned}
\left\| r^i - H(A^T A r^i + A^T e) \right\|_2 &= \left\| r_\Gamma^i + r_{\Gamma^c}^i - A_\Gamma^T A r^i + A_\Gamma^T e \right\|_2 \\
&\leq \left\| A_\Gamma^T A r^i - r_\Gamma^i \right\|_2 + \left\| r_{\Gamma^c}^i \right\|_2 + \left\| A_\Gamma^T e \right\|_2 \\
&\leq \left\| A_{\Gamma\cup\Omega}^T A r^i r_{\Gamma\cup\Omega}^i \right\|_2 + \left\| r_{\Gamma^c}^i \right\|_2 + \left\| A_\Gamma^T e \right\|_2 \\
&\leq \delta \left\| r^i \right\|_2 + \sqrt{1-\alpha_0^2} \left\| r^i \right\|_2 + \left(\frac{\beta_0}{\alpha_0} + \frac{\alpha_0 \beta_0}{\sqrt{1-\alpha_0^2}} + \sqrt{1+\delta} \right) \|e\|_2
\end{aligned}
$$

Proof of Geometric Convergence (continued)

$$\left\|x - x^{i+1}\right\|_2 = (1 + c_T)\left\|r^i - H(A^T A r^i + A^T e)\right\|_2$$

Combining RIP, lemma, and bound on $\left\|r^i - H(A^T A r^i + A^T e)\right\|_2$,

$$\left\|x - x^{i+1}\right\|_2 \leq \alpha\left\|x - x^i\right\|_2 + \beta\|e\|_2$$

where $\alpha = (1 + c_T)\left(\delta + \sqrt{1 - \alpha_0^2}\right)$ and

$$\beta = (1 + c_T)\left(\frac{\beta_0}{\alpha_0} + \frac{\alpha_0 \beta_0}{\sqrt{1 - \alpha_0^2}} + \sqrt{1 + \delta}\right)$$

This means AM-IHT exhibits robust signal recovery.

Geometric Converge in Noiseless Case

In the noiseless case,

$$\alpha = (1 + c_T)\left(\delta + \sqrt{1 - (c_H(1-\delta) - \delta)^2}\right)$$

For convergence, $\alpha < 1$. Assume δ is very small. In order for AM-IHT to converge,

$$\alpha \approx (1 + c_T)\left(\sqrt{1 - c_H^2}\right) < 1$$

$$c_H^2 > 1 - \frac{1}{(1 + c_T)^2}$$

AM-IHT achieves geometric convergence comparable to other model-based compressive sensing methods.

Approximate Model-IHT Algorithm

Algorithm 1 Approximate Model-IHT

function AM-IHT(y, A, t)
    $x^0 \leftarrow 0$
    for $i \leftarrow 0, \ldots, t$ **do**
        $b^i \leftarrow A^T(y - Ax^i)$
        $x^{i+1} \leftarrow T(x^i + H(b^i))$
    return x^{t+1}

Approximate Model Compressive Sampling Matching Pursuit

Approximate Model-CoSaMP

This algorithm for model-based compressive sensing with approximate projection oracles focuses on recovering signals from structured sparsity models.

Algorithm

Algorithm 2 Approximate Model-CoSaMP

function AM-CoSaMP(y, A, t)
  $x^0 \leftarrow 0$
  for $i \leftarrow 0, \ldots, t$ **do**
    $b^i \leftarrow A^T(y - Ax^i)$
    $\Gamma \leftarrow \mathrm{supp}(H(b^i))$
    $S \leftarrow \Gamma \cup \mathrm{supp}(x^i)$
    $z|_S \leftarrow A_S^{\dagger} y, \quad z|_{S^C} \leftarrow 0$
    $x^{i+1} \leftarrow T(z)$
  return x^{t+1}

Assumptions on Algorithm

1. $x \in \mathbb{R}^{\mathrm{n}}$ and $x \in \boldsymbol{\mathcal{M}}$
2. $y = Ax + e$ for $e \in \mathbb{R}^{\mathrm{m}}$
3. T is a $(c_T, \mathbb{M}, \mathbb{M}_T, 2)$-tail approximation oracle.
4. H is a $(c_H, \mathbb{M}_T \oplus \mathbb{M}, \mathbb{M}_H, 2)$-head approximation oracle.
5. A has the $(\delta, \mathbb{M} \oplus \mathbb{M}_T \oplus \mathbb{M}_H)$-model RIP.

Geometric convergence of AM-CoSaMP

Let $r^i = x - x^i$ where x^i is the signal estimate at iteration i.

$$\left\|r^{i+1}\right\|_2 \leq \alpha \left\|r^i\right\|_2 + \beta \|e\|_2$$

where $\alpha = (1 + c_T)\left(\sqrt{\frac{1+\delta}{1-\delta}}\sqrt{1 - \alpha_0^2}\right)$,

$$\beta = (1 + c_T)\left(\sqrt{\frac{1+\delta}{1-\delta}}\left(\frac{\beta_0}{\alpha_0} + \frac{\alpha_0\beta_0}{\sqrt{1-\alpha_0^2}}\right) + \frac{2}{\sqrt{1-\delta}}\right),$$

$\alpha_0 = c_H(1 - \delta) - \delta$ and $\beta_0 = (1 + c_H)\sqrt{1 + \delta}$

Geometric Convergence Proof

Using triangle equality, tail approximation, and RIP,

$$\begin{aligned}
\left\|r^{i+1}\right\|_2 &= \left\|x - x^{i+1}\right\|_2 \\
&\leq \left\|x^{i+1} - z\right\|_2 + \|x - z\|_2 \\
&\leq c_T\|x - z\|_2 + \|x - z\|_2 \\
&= (1 + c_T)\|x - z\|_2 \\
&\leq (1 + c_T)\frac{\|A(x-z)\|_2}{\sqrt{1-\delta}} \\
&= (1 + c_T)\frac{\|Ax-Az\|_2}{\sqrt{1-\delta}}
\end{aligned}$$

Geometric Convergence Proof (continued)

Because $Ax = y - e$ and $A_z = A_S z_S$,

$$\left\|r^{i+1}\right\|_2 \leq (1 + c_T)\left(\frac{\|y - A_S z_S\|_2}{\sqrt{1-\delta}} + \frac{\|e\|_2}{\sqrt{1-\delta}}\right)$$

$$\leq (1 + c_T)\left(\frac{\|y - A_S x_S\|_2}{\sqrt{1-\delta}} + \frac{\|e\|_2}{\sqrt{1-\delta}}\right)$$

Geometric Convergence Proof (continued)

$$y = Ax + e = A_S x_S + A_{S^c} x_{S^c} + e$$

$$\begin{aligned}
\left\| r^{i+1} \right\|_2 &\leq (1 + c_T) \frac{\left\| A_{S^c} x_{S^c} \right\|_2}{\sqrt{1-\delta}} + (1 + c_T) \frac{2\|e\|_2}{\sqrt{1-\delta}} \\
&\leq (1 + c_T) \frac{\sqrt{1+\delta}}{\sqrt{1-\delta}} \|x_{S^c}\|_2 + (1 + c_T) \frac{2\|e\|_2}{\sqrt{1-\delta}} \\
&= (1 + c_T) \sqrt{\frac{1+\delta}{1-\delta}} \left\| \left(x - x^i \right)_{S^c} \right\|_2 + (1 + c_T) \frac{2\|e\|_2}{\sqrt{1-\delta}} \\
&\leq \ (1 + c_T) \sqrt{\frac{1+\delta}{1-\delta}} \left\| r^i_{\Gamma^c} \right\|_2 + (1 + c_T) \frac{2\|e\|_2}{\sqrt{1-\delta}}
\end{aligned}$$

Geometric Convergence Proof (continued)

$$\left\|r^{i+1}\right\|_2 \leq \alpha\left\|r^i\right\|_2 + \beta\|e\|_2$$

where $\alpha = (1 + c_T)\left(\sqrt{\frac{1+\delta}{1-\delta}}\sqrt{1-\alpha_0^2}\right)$,

$$\beta = (1 + c_T)\left(\sqrt{\frac{1+\delta}{1-\delta}}\left(\frac{\beta_0}{\alpha_0} + \frac{\alpha_0\beta_0}{\sqrt{1-\alpha_0^2}}\right) + \frac{2}{\sqrt{1-\delta}}\right),$$

$\alpha_0 = c_H(1-\delta) - \delta$ and $\beta_0 = (1 + c_H)\sqrt{1+\delta}$

This means AM-CoSaMP exhibits robust signal recovery.

Geometric Converge in Noiseless Case

Assume $e = 0$ and δ is very small.
For convergence, $\alpha < 1$. Assume δ is very small. In order for AM-CoSaMP to converge,

$$\alpha \approx (1 + c_T)\left(\sqrt{1 - c_H^2}\right) < 1$$

$$c_H^2 > 1 - \frac{1}{(1 + c_T)^2}$$

This is identical to the convergence of AM-IHT.

Algorithm

Algorithm 2 Approximate Model-CoSaMP

function AM-CoSaMP(y, A, t)
  $x^0 \leftarrow 0$
  for $i \leftarrow 0, \ldots, t$ **do**
    $b^i \leftarrow A^T(y - Ax^i)$
    $\Gamma \leftarrow \mathrm{supp}(H(b^i))$
    $S \leftarrow \Gamma \cup \mathrm{supp}(x^i)$
    $z|_S \leftarrow A_S^{\dagger} y, \quad z|_{S^C} \leftarrow 0$
    $x^{i+1} \leftarrow T(z)$
  return x^{t+1}

Improved Recovery via Boosting

Why do these algorithms need to be boosted?

By definition, $c_T \geq 1$, and thus $c_H \geq \frac{\sqrt{3}}{2}$. If c_T is very large, this means the tail-approximation oracle can only give a crude approximation. This forces c_H to have to be very accurate, which can constrain the choice of approximation algorithms.

Boosting Algorithm

Algorithm 4 Boosting for head-approximation algorithms

function $\textsc{BoostHead}(x, H, t)$

  $\Omega_0 \leftarrow \{\}$

  for $i \leftarrow 1, \ldots, t$ **do**

    $\Lambda_i \leftarrow H(x_{[n] \setminus \Omega_{i-1}})$

    $\Omega_i \leftarrow \Omega_{i-1} \cup \Lambda_i$

  return Ω_t

Theorem

Let H be a $(c_H, \mathbb{M}, \mathbb{M}_H, p)$-head-approximation algorithm with $0 < c_H \leq 1$ and $p \geq 1$. Then, BoostHead(x, H, t) is a $\left(\left(1-\left(1-c_H^p\right)^t\right)^{\frac{1}{p}}, \mathbb{M}, \mathbb{M}_H, p\right)$-head-approximation algorithm. BoostHead runs in time $O(tT_H)$, where T_H is the time complexity of H.

Main Result

Let T and H be approximate projection oracles with $c_T \geq 1$ and $0 < c_H < 1$.

$$\gamma = \frac{\sqrt{1 - \left(\frac{1}{1 + c_T} - \delta\right)^2} + \delta}{1 - \delta}$$

$$t = \left\lceil \frac{\log(1 - \gamma^2)}{\log(1 - c_H^2)} \right\rceil + 1$$

If using AM-IHT with T and $\mathrm{BoostHead}(x, H, t)$ as projection oracles, the signal estimate will satisfy

$$\|x - \hat{x}\|_2 \leq C\|e\|_2$$

where $\hat{x}$ is the returned signal estimate.

Summary

Approximate Model-Based Algorithms

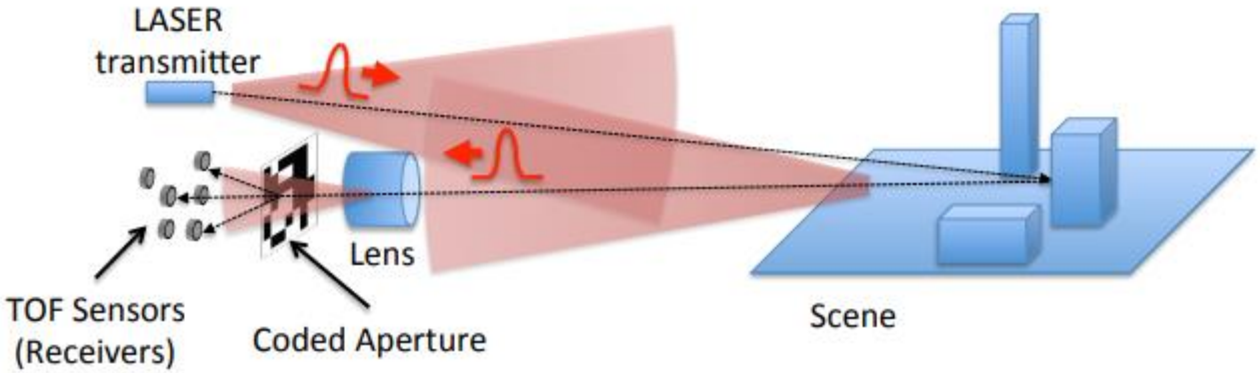
-AM-IHT

-AM-CoSaMP

Relaxation of Requirements via Boosting

Other research

Using model-based CoSaMP, the amount of LIDAR sensors needed can be reduced by as much as 85%.[2]



References

[1] Chinmay Hegde,Piotr Indyk,and Ludwig Schmidt. Approximation algorithms for model-based compressive sensing. Information Theory, IEEE Transactions, 61(9):5129–5147, 2015.

[2] A. Kadambi and P. T. Boufounos. Coded aperture compressive 3-d lidar. In 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pages 1166–1170, April 2015.