AMATH 840: Advanced Numerical Methods for Computational and Data Science

Winter 2023

Part 1: Compressive Sensing

1.2: Greedy and Thresholding-Based Methods

Winter 2023

Some Popular Algorithms for Compressive Sensing - Part 1

Problem: Given $A \in \mathbb{C}^{m \times n}$ and $\mathbf{y} \in \mathbb{C}^m$. Find

$$\mathbf{w} \in \mathbb{C}^n \quad \text{such that } \mathbf{y} = A\mathbf{w} \text{ and } \|w\|_0 \leq s.$$

We will go over three important algorithms:

- Orthogonal Matching Pursuit (OMP)
- Iterative Hard Thresholding (IHT)
- Hard Thresholding Pursuit (HTP)

We use the same numbers of Theorems, Lemmas, Propositions from "A Mathematical Introduction to Compressive Sensing", by S. Foucart and H. Rauhut.

Outline

Orthogonal Matching Pursuit (OMP)

Iterative Hard Thresholding (IHT)

Hard Thresholding Pursuit (HTP)

Orthogonal Matching Pursuit (OMP)

Problem: Given $A \in \mathbb{C}^{m \times n}$ and $\mathbf{y} \in \mathbb{C}^m$. Find

$$\mathbf{w} \in \mathbb{C}^n \quad \text{such that } \mathbf{y} = A\mathbf{w} \text{ and } \|w\|_0 \leq s.$$

Main Idea of OMP: Initialize $\mathcal{S}^0 = \emptyset \subseteq [n] := \{1, 2, \ldots, n\}$.

- Add one index to a target support $\mathcal{S}^k \subseteq [n]$ at each iteration, $\mathcal{S}^{k+1} = \mathcal{S}^k \cup \{j_{k+1}\}$, so that

$$\min_{\mathrm{supp}(z) \subseteq \mathcal{S}^{k+1}} \|\mathbf{y} - A\mathbf{z}\|_2 < \min_{\mathrm{supp}(z) \subseteq \mathcal{S}^{k}} \|\mathbf{y} - A\mathbf{z}\|_2.$$

- Update a target vector $\mathbf{w}^{k+1}$:

$$\mathbf{w}^{k+1} := \operatorname*{argmin}_{\mathrm{supp}(z) \subseteq \mathcal{S}^{k+1}} \|y - Az\|_2.$$

In the next few slides, we will derive the steps of OMP.

Orthogonal Matching Pursuit

Lemma 3.3. Given $A = [\mathbf{a}_1\, \mathbf{a}_2\, \ldots \mathbf{a}_n] \in \mathbb{C}^{m\times n}$ such that $\|\mathbf{a}_k\|_2 = 1,\ \forall k$. Given $S \subseteq [n]$ and $j \in [n]$. If

$$\mathbf{w} := \operatorname{argmin}\{\|\mathbf{y} - A\mathbf{z}\|_2 :\ \operatorname{supp}(\mathbf{z}) \subseteq S \cup \{j\}\},$$

then

$$\|\mathbf{y} - A\mathbf{w}\|_2^2 \leq \|\mathbf{y} - A\mathbf{v}\|_2^2 - |(A^*(\mathbf{y} - A\mathbf{v}))_j|^2,$$

for all $\mathbf{v} \in \mathbb{C}^n$ s.t. $\operatorname{supp}(\mathbf{v}) \subseteq S$.

$\Rightarrow$ OMP Algorithm: Choose $j = \arg\max_{j\in[n]} |(A^*(\mathbf{y} - A\mathbf{v}))_j|$.

Orthogonal Matching Pursuit

Proof Sketch.

Let $\mathbf{v} \in \mathbb{C}^n$ s.t. $\mathrm{supp}(\mathbf{v}) \subseteq S$.

- For any $t \in \mathbb{C}$, $\mathrm{supp}(\mathbf{v} + t\mathbf{e}_j) \subseteq S \cup \{j\}$, show that

$$\|\mathbf{y} - A(\mathbf{v} + t\mathbf{e}_j)\|_2^2 \geq \|\mathbf{y} - A\mathbf{v}\|_2^2 - |(A^*(\mathbf{y} - A\mathbf{v}))_j|^2,$$

  and the quality holds when $|t| = |(A^*(\mathbf{y} - A\mathbf{v}))_j|$.

- Therefore,

$$\begin{aligned}\|\mathbf{y} - A\mathbf{w}\|_2^2 &= \min\{\|\mathbf{y} - A\mathbf{z}\|_2^2 : \ \mathrm{supp}(\mathbf{z}) \subseteq S \cup \{j\}\} \\ &\leq \min_{t \in \mathbb{C}} \|\mathbf{y} - A(\mathbf{v} + t\mathbf{e}_j)\|_2^2 \\ &\leq \|\mathbf{y} - A\mathbf{v}\|_2^2 - |(A^*(\mathbf{y} - A\mathbf{v}))_j|^2.\end{aligned}$$

$\square$

Proof. We have

$$\begin{aligned}\|y - A(v + te_j)\|_2^2 &= \|y - Av - tAe_j\|_2^2 \\ &= \|y - Av\|_2^2 + |t|^2 \|Ae_j\|_2^2 - 2\text{Re}\langle y - Av, tAe_j\rangle \\ &= \|y - Av\|_2^2 + |t|^2 - 2\text{Re}\left(\bar{t}e_j^T A^*(y - Av)\right) \\ &\geq \|y - Av\|_2^2 + |t|^2 - 2|t||(A^*(y - Av))_j| \\ &\geq \|y - Av\|_2^2 - |(A^*(y - Av))_j|^2.\end{aligned}$$

The equality holds when $|t| = |(A^*(y - Av))_j|$ and

$$\text{Re}\left(\bar{t}e_j^T A^*(y - Av)\right) = |t|\,|\bar{t}e_j^T A^*(y - Av)|.$$

Therefore,

$$\min_{t\in\mathbb{C}} \|y - A(v + te_j)\|_2^2 = \|y - Av\|_2^2 - |(A^*(y - Av))_j|^2,$$

and

$$\begin{aligned}\|y - Aw\|_2^2 = \min_{\text{supp}(z)\subseteq S\cup\{j\}} \{\|y - Az\|_2^2\} &\leq \min_{t\in\mathbb{C}} \|y - A(v + te_j)\|_2^2 \\ &\leq \|y - Av\|_2^2 - |(A^*(y - Av))_j|^2.\end{aligned}$$

Orthogonal Matching Pursuit

For $z \in \mathbb{C}^n$ and $S \subseteq [n]$, denote $v_S := v|_S \in \mathbb{C}^{|S|}$, the restriction of v onto the index set S.

Lemma 3.4. Given an index set $S \subseteq [n]$. If

$$\mathbf{v} := \operatorname*{argmin}_{\mathbf{z} \in \mathbb{C}^n} \{\|\mathbf{y} - A\mathbf{z}\|_2, \operatorname{supp}(\mathbf{z}) \subseteq S\},$$

then

$$A_S^* \mathbf{y} = A_S^* A_S \mathbf{v}_S, \quad \text{i.e.,} \quad (A^*(\mathbf{y} - A\mathbf{v}))_S = 0.$$

⇒ OMP algorithm:

$$\mathbf{w}^{k+1} := \operatorname*{argmin}_{\operatorname{supp}(\mathbf{z}) \subseteq \mathcal{S}^{k+1}} \|\mathbf{y} - A\mathbf{z}\|_2$$

$$\mathbf{w}^{k+1}_{\mathcal{S}^{k+1}} = A^{\dagger}_{\mathcal{S}^{k+1}} \mathbf{y}.$$

Proof. We rewrite the constraint optimization problem

$$\mathbf{v} := \underset{\mathbf{z}\in\mathbb{C}^n}{\operatorname{argmin}}\{\|\mathbf{y} - A\mathbf{z}\|_2, \operatorname{supp}(\mathbf{z}) \subseteq S\}$$

to an unconstraint optimization problem:

$$v = \underset{\mathbf{z}\in\mathbb{C}^n}{\operatorname{argmin}}\{\|\mathbf{y} - A_S\mathbf{z}_S\|_2\} = \underset{\mathbf{u}\in\mathbb{C}^{|S|}}{\operatorname{argmin}}\{\|\mathbf{y} - A_S\mathbf{u}\|_2\}$$

The least square solution is

$$A_S^*\mathbf{y} = A_S^*A_S\mathbf{v}_S.$$

Orthogonal Matching Pursuit: Pseudocode

Orthogonal Matching Pursuit

Input: Measurement matrix $A \in \mathbb{C}^{m \times n}$ with ℓ_2-normalized columns, measurement vector $\mathbf{y} \in \mathbb{C}^m$, sparsity level s or tolerance ε.

Initialization: $\mathcal{S}^0 = \emptyset, \mathbf{w}^0 = 0$.

Iteration: Repeat until Stopping Criterion is met.

$$j_{k+1} := \arg\max_{j \in [n]} \{ |(A^*(\mathbf{y} - A\mathbf{w}^k))_j| \}$$

$$\mathcal{S}^{k+1} := \mathcal{S}^k \cup \{j_{k+1}\}$$

$$\text{Find } \mathbf{w}^{k+1} \; s.t \; A^*_{\mathcal{S}^{k+1}} \mathbf{y} = A^*_{\mathcal{S}^{k+1}} A_{\mathcal{S}^{k+1}} \mathbf{w}^{k+1}_{\mathcal{S}^{k+1}}$$

Output: The sparse vector $\mathbf{w}^{\#}$.

OMP

```
import numpy as np
from sklearn.linear_model import OrthogonalMatchingPursuit as omp
```

Remarks about ℓ_2-normalized columns of the measurement matrix A:

- It is not strictly required. However, if we work with a general matrix A, we need to adjust the steps in the iterations accordingly.
- For $A = [\mathbf{a}_1 \, \mathbf{a}_2 \, \ldots \mathbf{a}_n] \in \mathbb{C}^{m \times n}$ and $\mathbf{w} = [w_1, w_2, \ldots, w_n]^T \in \mathbb{C}^n$, we have

$$A\mathbf{w} = \sum_{k=1}^{n} w_k \mathbf{a}_k = \sum_{k=1}^{n} \left(w_k \|\mathbf{a}_k\|_2 \right) \frac{\mathbf{a}_k}{\|\mathbf{a}_k\|_2} = \hat{A}\hat{\mathbf{w}},$$

  where $\hat{A} = \begin{bmatrix} \frac{\mathbf{a}_1}{\|\mathbf{a}_1\|_2} & \frac{\mathbf{a}_2}{\|\mathbf{a}_2\|_2} & \cdots & \frac{\mathbf{a}_n}{\|\mathbf{a}_n\|_2} \end{bmatrix}$ and $\hat{w}_k = w_k \|\mathbf{a}_k\|_2$.

  Using OMP to solve:

$$\text{Find } \hat{\mathbf{w}} \text{ sparse s.t. } \mathbf{y} = \hat{A}\hat{\mathbf{w}},$$

  and get back to the original solution $\mathbf{w}$.

Orthogonal Matching Pursuit: Some Remarks

1. Recall: $j_{k+1} := \arg\max_{j \in [n]} \{ |(A^*(y - Aw^k))_j| \}, \quad A^*_{S^k} y = A^*_{S^k} A_{S^k} w^k_{S^k}.$

2. How to speed up?

   - Use QR decomposition of the matrix A_{S^k}
   - Fast algorithm to update QR decomposition of $A_{S^{k+1}}$ from A_{S^k}.
   - Use fast matrix-vector multiplication for A (Fourier transform)

3. Stopping criterions:

   - If we know the sparsity level s, stop after Cs iterations.
   - To account for measurement and computational errors, stop when

$$\|y - Aw^k\|_2 \leq \varepsilon, \quad \text{or} \quad \|A^*(y - Aw^k)\|_\infty \leq \varepsilon,$$

   where ε is a chosen tolerance.

Outline

Orthogonal Matching Pursuit (OMP)

Iterative Hard Thresholding (IHT)

Hard Thresholding Pursuit (HTP)

Iterative Hard Thresholding (IHT)

Problem: Given $A \in \mathbb{C}^{m\times n}$ and $\mathbf{y} \in \mathbb{C}^m$. Find

$$\mathbf{w} \in \mathbb{C}^n \quad \text{such that } \mathbf{y} = A\mathbf{w} \text{ and } \|w\|_0 \leq s.$$

Motivation: Solve $A\mathbf{w} = \mathbf{y}$ by iteration methods. Indeed, from $\mathbf{y} = A\mathbf{w}$, we have:

$$\begin{aligned} 0 &= A^*(\mathbf{y} - A\mathbf{w}) \\ \mathbf{w} &= \mathbf{w} + \gamma A^*(\mathbf{y} - A\mathbf{w}), \quad \text{where} \quad \gamma \in \mathbb{R} \\ \mathbf{w}^{k+1} &= \mathbf{w}^k + \gamma A^*(\mathbf{y} - A\mathbf{w}^k) \quad \text{(fixed-point iterate)} \end{aligned}$$

Then we apply a projection method to ensure the iterates $\mathbf{w}^k$ have at most s nonzero entries.

Banach Fixed-Point Theorem

Theorem

Given $p \in [1, \infty]$. *Assume that:*

(i) M *is a closed, nonempty set in the Banach space* $\mathbb{C}^n$ *or* $\mathbb{R}^n$.

(ii) *The operator* $\mathcal{A} : M \to M$ *is L-contractive for some fixed* $L \in [0, 1)$*:*

$$\|\mathcal{A}\mathbf{u} - \mathcal{A}\mathbf{v}\|_p \leq L\|\mathbf{u} - \mathbf{v}\|_p \quad \text{for all } \mathbf{u}, \mathbf{v} \in M.$$

Then the following statements hold true:

1. *The equation* $\mathbf{u} = \mathcal{A}\mathbf{u}, \mathbf{u} \in M$, *has exactly one solution* $\mathbf{u}_* \in M$.
2. *For each given* $\mathbf{u}^0 \in M$, *the sequence* $\{\mathbf{u}^k\}$ *constructed by the iteration method,* $\mathbf{u}^{k+1} = \mathcal{A}\mathbf{u}^k$, *converges to the unique solution* $\mathbf{u}_*$ *of* $\mathbf{u} = \mathcal{A}\mathbf{u}$ *and*

$$\|\mathbf{u}^k - \mathbf{u}_*\|_p \leq \frac{L^k}{1 - L}\|\mathbf{u}^1 - \mathbf{u}^0\|_p,$$

Iterative Hard Thresholding (IHT)

Denote

$$L_s(\mathbf{z}) := \text{Index set of } s \text{ largest absolute entries of } \mathbf{z} \in \mathbb{C}^n,$$
$$H_s(\mathbf{z}) := \mathbf{z} L_s(\mathbf{z}).$$

Iterative Hard Thresholding (IHT)

Input: Measurement matrix A, measurement vector $\mathbf{y}$, sparsity level s, hyperparameter γ.

Initialization: s-sparse vector $\mathbf{w}^0$, typically $\mathbf{w}^0 = 0$.

Iteration: Repeat until a stopping criterion is met.

$$\mathbf{w}^{k+1} = H_s(\mathbf{w}^k + \gamma A^*(\mathbf{y} - A\mathbf{w}^k)).$$

Output: The s-sparse vector $\mathbf{w}^{\#}$.

Remark: The IHT does not require the computation of any orthogonal projection.

Outline

Orthogonal Matching Pursuit (OMP)

Iterative Hard Thresholding (IHT)

Hard Thresholding Pursuit (HTP)

Hard Thresholding Pursuit (HTP)

Motivation: Basic Thresholding + Orthogonal Projection

Hard Thresholding Pursuit

Input: Measurement matrix A, measurement vector $\mathbf{y}$, sparsity level s, hyperparameter γ.

Initialization: s-sparse vector $\mathbf{w}^0$, typically $\mathbf{w}^0 = 0$.

Iteration: Repeat until a stopping criterion is met.

$$S^{k+1} = L_s(\mathbf{w}^k + \gamma A^*(\mathbf{y} - A\mathbf{w}^k)),$$
$$\mathbf{w}^{k+1} = \underset{\mathbf{z}\in\mathbb{C}^n}{\text{argmin}}\{\|\mathbf{y} - A\mathbf{z}\|_2, \quad \mathrm{supp}(\mathbf{z}) \subseteq S^{k+1}\}$$

Output: The s-sparse vector $\mathbf{w}^\#$.

Remarks:

- If we don't know the sparsity s, we can base on the absolute values of entries in each iteration and decide how many entries we would like to keep. For example, in the input of IHT, instead of using the sparsity level s, we can choose a hyperparameter ε and H_s is replaced by H_ε, where

$$H_\varepsilon(z) = \begin{cases} z & \text{if } |z| > \varepsilon, \\ 0 & \text{otherwise} \end{cases}$$

- The error estimations rely on the property of the matrix A (for e.g, RIP-like condition).