E9 203 Compressed Sensing & Sparse SP

12 Mar. 2021.

Last time:

- ℓ_1 recovery $\Rightarrow$ m-sparse solns.
- $\ell_1 \leftrightarrow$ LP, QCBP in the complex case
- Related convex opt. pbs.:
  - QCBP
  - BPDN
  - LASSO
  - Dantzig selector
- Orthogonal matching pursuit. (OMP).

Today:

- OMP when cols of A are not normalized
- Properties of OMP
- Thresholding based methods

OMP when cols of A are not normalized:

Init: $k=0$, $x^{(0)}=0$, $r^{(0)} = y - Ax^{(0)}$, $S^{(0)} = \{\emptyset\}$

Repeat: $k \leftarrow k+1$

(a) Find $\varepsilon_j = \min_{z_j \in \mathbb{R}} \| a_j z_j - r^{(k-1)} \|_2^2$, $j \in [N]$ [Compute errors]

Optimal $z_j^* = \dfrac{a_j^T r^{(k-1)}}{\|a_j\|_2^2}$

(b) Find $j_0 = \min_{j \in [N]} \varepsilon_j$

Update $S^{(k)} = S^{(k-1)} \cup \{j_0\}$

(c) Update $x^{(k)} = \arg\min_{x \in \mathbb{R}^N, \, supp(x) \subseteq S^{(k)}} \|Ax - y\|_2^2$

(d) update $r^{(k)} = y - Ax^{(k)}$.

Until: stopping criterion, e.g., $\|r^{(k)}\|_2 \le \varepsilon_{min}$.

Note: in (a)

$$\varepsilon_j = \min_{z_j} \| a_j z_j - r^{(k-1)} \|_2^2, \quad z_j^* = \frac{a_j^T r^{(k-1)}}{\|a_j\|_2^2}$$

$$= \left\| \frac{a_j^T r^{(k-1)}}{\|a_j\|_2^2} a_j - r^{(k-1)} \right\|_2^2$$

$$= \|r^{(k-1)}\|_2^2 - \underbrace{\frac{(a_j^T r^{(k-1)})^2}{\|a_j\|_2^2}}_{\text{only term that depends on } j.}$$

$\Rightarrow$ Finding the smallest ε_j is equivalent to finding the normalized column that is most aligned with the residual. $\left| \left(\frac{a_j}{\|a_j\|}\right)^T r^{(k-1)} \right|^2$

Lemma 3.3

$A \in \mathbb{C}^{m \times N}$, ℓ_2 normalized cols. Given $S \subseteq [N]$, v supported on S, $j \in [N]$, if $\boxed{w \triangleq \arg\min_{z \in \mathbb{C}^N} \{ \|y - Az\|_2, \, supp(z) \subseteq S \cup \{j\} \}}$

then $\boxed{\|y - Aw\|_2^2 \le \|y - Av\|_2^2 - |(A^H(y - Av))_j|^2}$ $\Leftarrow$

Remark 1: $\|y - Aw\|_2^2 \le \|y - Av\|_2^2$ not surprising. The lemma tells us that adding col j to the support reduces the norm of the residual even more.

Remark 2: v can be any vec supported on S, in particular, it could be the "best" vec. supported on S:

$$\arg\min_{z \in \mathbb{C}^N} \{ \|y - Az\|_2, \, supp(z) \subseteq S \}.$$

Proof: $supp(v + te_j) \subseteq S \cup \{j\}$

$\Rightarrow \|y - Aw\|_2^2 \le \min_{t \in \mathbb{C}} \|y - A(v + te_j)\|_2^2$. $\Leftarrow$

Polar form: $t = \rho e^{j\theta}$, $\rho \ge 0$, $\theta \in [0, 2\pi) \Rightarrow$

$$\|y - A(v + te_j)\|_2^2 = \|y - Av - tAe_j\|_2^2$$

$$= \|y - Av\|_2^2 + \underbrace{|t|^2}_{\rho^2} \underbrace{\|Ae_j\|_2^2}_{(=1 \because A \text{ has } \ell_2\text{-normalized cols.})} - 2\,\mathrm{Re}\{ \bar{t}\, \underbrace{e_j^H A^H (y - Av)}_{[A^H(y - Av)]_j} \}$$

$$\ge \|y - Av\|_2^2 + \underbrace{\rho^2 - 2\rho \boxed{|(A^H(y - Av))_j|}}$$

$\uparrow$ equality when $\theta = \angle (A^H(y - Av))_j$ — minimized when $\rho = |(A^H(y - Av))_j|$

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

$\uparrow$ $\because$ Equality can be achieved by choosing the optimal t.

Of course, w could reduce the ℓ_2 norm of the residual even further, so the result follows. $\square$

[HW: Review properties of the pseudoinverse, Sec. A2]

Stopping criterion for OMP:

1. $Ax^{(n)} = y$. (Doesn't work if there are meas./comput. errors)
2. $\|y - Ax^{(n)}\|_2 \le \varepsilon$ } $\varepsilon > 0$: tolerance threshold
3. $\|A^H(y - Ax^{(n)})\|_\infty \le \varepsilon$
4. If sparsity s is known, (or an est. is available), can use $n = s$. $\Rightarrow$ Results in an s-sparse output.

When will OMP succeed?

Prop. 3.5 Given $A \in \mathbb{C}^{m \times N}$, normalized cols. Every $0 \ne x \in \mathbb{C}^N$, $supp(x) \subseteq S$, $|S| = s$, is recovered from $y = Ax$ after at most s iterations of OMP iff

(i) A_S is injective

(ii) $\max_{j \in S} |(A^H r)_j| > \max_{\ell \in \bar{S}} |(A^H r)_\ell|$ (Not $\forall S$, $|S| \le s$)

for all $0 \ne r \in \{ Az, \, supp(z) \subseteq S \}$.

Proof: Necessary $\boxed{\Rightarrow}$

Assume OMP recovers all vecs. supported on S in at most $s = |S|$ iterations.

$\Rightarrow A_S$ is injective $\because$ 2 vecs supported on S with the same measurement vec. must be equal by the hypothesis.

In OMP, the index chosen in the 1st iteration stays in the support set of the output. $\Rightarrow$ If $y = Ax$, $x \in \mathbb{C}^N$, and x is exactly supported on S, then $\ell \in \bar{S}$ cannot be chosen in the 1st iteration, i.e.,

$$\max_{j \in S} |(A^H y)_j| > |(A^H y)_\ell| \quad \forall \ell \in \bar{S}$$

$$\Rightarrow \max_{j \in S} |(A^H y)_j| > \max_{\ell \in \bar{S}} |(A^H y)_\ell|.$$

This is true $\forall y \in \{ Az, \, supp(z) \subseteq S \}$.

Thus, the two conditions are necessary.

Sufficient $\boxed{\Leftarrow}$

Assume $Ax^{(1)} \ne y$, $Ax^{(2)} \ne y$, $\ldots$, $Ax^{(s-1)} \ne y$ WLOG

Given $0 \le n \le s-1$, if $S^n \subseteq S$, then

$r^{(n)} = y - Ax^{(n)} \in \{ Az, \, supp(z) \subseteq S \}$

$\Rightarrow j_{n+1} \in S$ by (ii) $\left[\text{OMP picks } j_{n+1} = \underset{j\in[N]}{\arg\max} \{ |(A^H r^{(n)})_j| \}\right]$

$\Rightarrow S^{n+1} = S^n \cup \{j_{n+1}\} \subseteq S$

Hence, $S^n \subseteq S \;\forall\; 0 \le n \le s.$

Now, given $1 \le n \le s-1$, $(A^H r^{(n)})_{S^n} = 0$ by the orthogonality principle $\Rightarrow j_{n+1} \notin S^n$

(Else, $A^H r^{(n)} = 0 \Rightarrow r^{(n)} = 0$, contradicting $Ax^{(1)} \neq y, \dots, Ax^{(s-1)} \neq y$.)

Thus, $|S^n| = n$, i.e., $S^s = S$.

Hence $Ax^{(s)} = y$ $\because$ OMP finds $x^{(s)} = \underset{z\in\mathbb{C}^N}{\arg\min} \left\{ \begin{array}{l} \|y - Az\|_2 \\ \text{s.t. } \operatorname{supp}(z) \subseteq S^s = S \end{array} \right\}$

By the injectivity of A_S, the above $\Rightarrow x^{(s)} = x$, which completes the proof. $\square$

Notation: $z \in \mathbb{C}^N$.

$L_s(z) \triangleq$ Index set of s largest absolute entries of z.

$H_s(z) \triangleq z_{L_s(z)}$ (Best s-term approx of z).

Thresholding based methods

Basic Thresholding: (BT)

$$\boxed{S^\# = L_s(A^H y)}$$

$$x^\# = \underset{z\in\mathbb{C}^N}{\arg\min} \{ \|y - Az\|_2,\ \operatorname{supp}(z) \subseteq S^\# \}$$

output $x^\#$.

Prop. 3.7 $x \in \mathbb{C}^N$, $S \triangleq \operatorname{supp}(x)$ is recovered from $y = Ax$ via BT iff

$$\left[\min_{j\in S} |(A^H y)_j| > \max_{l\in \bar{S}} |(A^H y)_l| . \right].$$