Chapter 5. Mutual Incoherence

I-Liang Chern

October 12, 2016

Goal of this chapter.

- In compressive sensing, the analysis of recovery algorithms usually involves a quantity that measures the suitability of the measurement matrix.
- The coherence is a very simple such measure of quality. In general, the smaller the coherence, the better the recovery algorithms perform.
- Goal: Introduce the concept of coherence and give sufficient conditions expressed in terms of the coherence that guarantee the success of orthogonal matching pursuit, basis pursuit, and thresholding algorithms. [1]

[1] This is a note from S. Foucart and H. Rauhut, A Mathematical Introduction to Compressive Sensing, Springer 2013.

What's kind of measurement matrix we like

- Let $\mathbf{A} = [a_1, \cdots, a_N]$, N column vectors in $\mathbb{C}^m$. We normalize them by $\|a_j\|_2 = 1$. Let $A^* = [b_1, \cdots, b_m]$, row vectors in C^N.
- Usually, $m << N$, this means $\{a_1, \cdots, a_N\}$ a redundant set in $\mathbb{C}^m$, while $\{b_1, \cdots, b_m\} \subset \mathbb{C}^N$ will not be enough to span $\mathbb{C}^N$.
- We want $\{b_i\}$ less correlated, the best case is that they are orthogonal.
- We want each a_j equally important in representing $\mathbf{y}$. The best is that they are equiangular. This means that

$$|\langle a_i, a_j \rangle| = c \text{ for all } i \neq j.$$

Motivation to define incoherence

Theorem

Given $A \in \mathbb{C}^{m \times N}$, the following properties are equivalent:

(a) *Every s-sparse vector $x \in \mathbb{C}^N$ is the unique s-sparse solution of $Az = Ax$, that is, if $Ax = Az$ and $x, z \in \Sigma_s$, then $x = z$.*

(b) *The null space $ker A$ does not contain any $2s$-sparse vector other than the zero vector,that is, $ker A \cap \Sigma_{2s} = \{0\}$.*

(c) *For every $S \subset [N]$ with $|S| \leq 2s$, the submatrix A_S is injective as a map from $\mathbb{C}^S$ to $\mathbb{C}^m$.*

(d) *Every set of $2s$ columns of A is linearly independent.*

Motivation to define incoherence

(e) $A_S^* A_S : \mathbb{C}^S \to \mathbb{C}^S$ is invertible;

(f) That is, the matrix $(\langle a_i, a_j \rangle)_{i,j \in S}$ is invertible;

Measure the coherence

Let $\mathbf{A} = [a_1, ..., a_N]$ be an $m \times N$ matrix with $\|a_j\|_2 = 1 \ \forall j$.

Definition

1. Coherence of $\mathbf{A}$ is defined to be

$$\mu(\mathbf{A}) = \max_{i \neq j} |\langle a_i, a_j \rangle|.$$

2. The ℓ_1-coherence function: for $1 \leq s \leq N-1$

$$\mu_1(s) := \max_{i \in [N]} \max \{ \sum_{j \in S} |\langle a_i, a_j \rangle|, S \subset [N], |S| = s, i \notin S \}$$

Question: How small of μ or $\mu_1(s)$ leads to (P1) $\Leftrightarrow$ (P0)?

Simple properties

- $\mu \leq 1$
- $\mu \leq \mu_1(s) \leq s\mu$
- $\max\{\mu_1(s), \mu_1(t)\} \leq \mu_1(s+t) \leq \mu_1(s) + \mu_1(t)$

Proof.

1. Because $|\langle a_i, a_j \rangle| \leq \|a_i\| \|a_j\| = 1$
2. By definition $\mu(A) = \mu_1(1) \leq \mu_1(s)$ for $s \geq 1$.

$$\mu_1(s) \leq \max_{i \in [N]} \max_{|S|=s, i \notin S} \sum_{j \in S} |\langle a_i, a_j \rangle| \leq s\mu$$

3.

$$\begin{aligned}
\mu_1(s+t) &= \max_i \max_{|S \cup T|=s+t, i \notin S \cup T} \sum_{j \in S \cup T} |\langle a_i, a_j \rangle| \\
&= \max_i \left(\max_{|S|=s, i \notin S} \sum_{j \in S} + \max_{|T|=t, i \notin T} \sum_{j \in T} \right) |\langle a_i, a_j \rangle| \\
&\leq \max_i \max_{|S|=s, i \notin S} \sum_{j \in S} |\langle a_i, a_j \rangle| + \max_i \max_{|T|=t, i \notin T} \sum_{j \in T} |\langle a_i, a_j \rangle| = \mu_1(s) + \mu_1(t)
\end{aligned}$$

Theorem

We have: for all s-sparse vector x

$$(1-\mu_1(s-1))\|x\|_2^2 \le \|Ax\|_2^2 \le (1+\mu_1(s-1))\|x\|_2^2.$$

Equivalently, the spectrum

$$\sigma(A_S^* A_S) \subset [1-\mu_1(s-1), 1+\mu_1(s-1)]$$

for all S with $|S| \le s$. In particular, $A_S^* A_S$ is invertible for all $|S| \le s$ if

$$\mu_1(s-1) < 1.$$

Proof

1. For any x with support in S, we have

$$\|Ax\|^2 = \langle A_S x_S, A_S x_S \rangle = \langle A_S^* A_S x_S, x_S \rangle,$$

   which is bounded above and below by λ_{max} and λ_{min} of $A_S^* A_S$.
2. The diagonal entries of $A_S^* A_S$ are $\|a_j\|^2 = 1$. By Gershgorin's disk theorem, the eigenvalues of $A_S^* A_S$ lie in the union of disks centered at 1 with radii

$$r_j = \sum_{\ell \in S, \ell \neq j} |\langle a_j, a_\ell \rangle| \leq \mu_1(s-1), \quad j \in S.$$

   Since all eigenvalues are real ($A_S^* A_S$ self-adjoint), they lie in $[1-\mu_1(s-1), 1+\mu_1(s-1)]$.

Corollary

If $\mu_1(s) + \mu_1(s-1) < 1$*, then for any* S *with* $|S| \leq 2s$*, the matrix* $A_S^* A_S$ *is invertible and* A_S *is injective. In particular, the conclusion holds if*

$$\mu < \frac{1}{2s-1}.$$

Proof.

1. $\mu_1(2s-1) < \mu_1(s) + \mu_1(s-1) < 1$. Thus, $A_S^* A_S$ is invertible for S with $|S| \leq 2s$.
2. $\mu_1(s) + \mu_1(s-1) \leq (2s-1)\mu < 1$. Thus, if $\mu < 1/(2s-1)$, then $\mu_1(s) + \mu_1(s-1) < 1$.

Theorem

If $\mu_1(s) + \mu_1(s-1) < 1$*, then every* s*-sparse vector* x *can be recovered from the measurement* $y = Ax$ *via basis pursuit.*

Proof.

1. We want to show $\|v_S\|_1 < \|v_{\bar{S}}\|_1$ for $v \in N(A)$.
2. From $Av = 0$, we have $\sum_j a_j v_j = 0$. Taking inner product with a_i, with $i \in S$, we get

$$v_i \langle a_i, a_i \rangle = -\sum_{j \neq i} v_j \langle a_j, a_i \rangle$$

$$|v_i| \leq \sum_{\ell \in \bar{S}} |v_\ell| |\langle a_\ell, a_i \rangle| + \sum_{j \in S, j \neq i} |v_j| |\langle a_j, a_i \rangle|$$

$$\begin{aligned} \|v_S\|_1 &\leq \sum_{\ell \in \bar{S}} |v_\ell| \sum_{i \in S} |\langle a_\ell, a_i \rangle| + \sum_{j \in S} |v_j| \sum_{i \in S, i \neq j} |\langle a_j, a_i \rangle| \\ &\leq \mu_1(s) \|v_{\bar{S}}\|_1 + \mu_1(s-1) \|v_S\|_1 \end{aligned}$$

3. $\mu_1(s)\|v_S\|_1 < (1 - \mu_1(s-1))\|v_S\|_1 \leq \mu_1(s)\|v_{\bar{S}}\|_1$.

Theorem

If $\mu_1(s) + \mu_1(s-1) < 1$, then every s-sparse vector x can be recovered from the measurement $y = Ax$ after at most s iterations of orthogonal matching pursuit.

1. We want to show that for any $S \subset [N]$ with $|S| \leq s$, $\max_{j \in S} |\langle r, a_j \rangle| > \max_{\ell \in \bar{S}} |\langle r, a_\ell \rangle|$ for all $r = \sum_{i \in S} r_i a_i$.
2. Let $|r_k| = \max_{i \in S} |r_i| > 0$. For $\ell \in \bar{S}$,

$$|\langle r, a_\ell \rangle| = \left| \langle \sum_{i \in S} r_i a_i, a_\ell \rangle \right| \leq \sum_{i \in S} |r_i| |\langle a_i, a_\ell \rangle| \leq |r_k| \mu_1(s).$$

$$|\langle r, a_k \rangle| \geq |r_k| |\langle a_k, a_k \rangle| - \sum_{i \in S, i \neq k} |r_i| |\langle a_i, a_k \rangle| \geq |r_k| (1 - \mu_1(s-1))$$

3. If $1 - \mu_1(s-1) > \mu_1(s)$, then $|\langle r, a_k \rangle| > |\langle r, a_\ell \rangle|$ for $\ell \in \bar{S}$.
4. A_S is injective from $\mu_1(s) + \mu_1(s-1) < 1$ (Corollary).

Theorem

Suppose supp $x = S$ *and* $|S| = s$. *If*

$$\mu_1(s) + \mu_1(s-1) < \frac{\min_{i \in S} |x_i|}{\max_{i \in S} |x_i|}$$

then x *can be recovered from the measurement* $y = Ax$ *via basic thresholding.*

Theorem

If

$$2\mu_1(s) + \mu_1(s-1) < 1,$$

then every s*-sparse vector* x *is exactly recovered from the measurement vector* $y = Ax$ *after at most* s *iterations of hard thresholding pursuit.*

Matrices with small coherence: equiangular tight frame

Proposition

Given $\{a_1, ..., a_N\}$ *in* $\mathbb{C}^m$*, the following three are equivalent*

(a) $\exists \lambda > 0$ *such that for any* $x \in \mathbb{C}^m$*,* $x = \lambda \sum_{j=1}^{N} \langle x, a_j \rangle a_j$

(b) $AA^* = \frac{1}{\lambda} Id_m$ *for some* $\lambda > 0$

(c) *For any* $x \in \mathbb{C}^m$*,* $\|x\|_2^2 = \lambda \sum_{j=1}^{N} |\langle x, a_j \rangle|^2$.

Remark.

- Such $\{a_1, ..., a_N\}$ is called a tight frame.
- The ideal measurement matrix should be an equiangular tight frame.

Proof.

1. (a)⇔(b): for any $x \in \mathbb{C}^m$,

$$AA^*x = \sum_j a_j a_j^* x = \sum_j \langle x, a_j \rangle a_j = \frac{1}{\lambda} x$$

2. (a)⇒(c):

$$\langle x, x \rangle = \left\langle \lambda \sum_j \langle x, a_j \rangle a_j, x \right\rangle = \lambda \sum_j |\langle x, a_j \rangle|^2$$

3. (c)⇒(a): Use polarization identity

$$\langle x, y \rangle = \frac{1}{4} \left[\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2 \right]$$

we can obtain

$$\langle x, y \rangle = \lambda \sum_j \langle x, a_j \rangle \langle a_j, y \rangle = \left\langle \lambda \sum_j \langle x, a_j \rangle a_j, y \right\rangle .$$

Equiangular tight frame

Definition

A family of normalized vectors $\{a_1, \ldots, a_N\}$ is called equiangular if

$$|\langle a_i, a_j \rangle| = c \text{ for all } i \neq j.$$

Examples

- In $\mathbb{R}^2$, let $a_1 = e_1$, $a_2 = (-1, \sqrt{3})/2$, $a_3 = (-1, -\sqrt{3})/2$ form an ETF.
- The windowed Fourier basis (Garbor) forms an ETF.
- Splines
- Wavelets, Framelets

Matrices with small coherence

Theorem

- *Let* A *be* $m \times N$ *matrix with normalized column vectors.*

Then

$$\mu_1(s) \geq s\sqrt{\frac{N-m}{m(N-1)}} \quad \textit{whenever } s < \sqrt{N-1}.$$

The equality holds $\Leftrightarrow$ $\{a_1, ..., a_N\}$ *are equiangular tight frame.*

- *For equiangular system* $\{a_1, ..., a_N\}$ *in* $\mathbb{C}^m$*, it holds*

$$N \leq m^2.$$

The equality holds $\Leftrightarrow$ $\{a_1, ..., a_N\}$ *is also a tight frame.*

Minimal number of measurements

For the matrices with smallest coherence: then

- $N = m^2$
- $\mu_1(s) = s\sqrt{\frac{N-m}{m(N-1)}} \sim s\frac{c}{\sqrt{m}}$
- The condition $\mu_1(s) + \mu_1(s-1) < 1$ gives

$$(2s-1)\frac{c}{\sqrt{m}} < 1$$

- This requires $m \geq Cs^2$.
- We will see later that exact recovery can be obtained for

$$m \geq C\delta^{-2} s \ln(eNs)$$

  with $\delta_s \leq \delta$, where δ_s is another "coherence" measurement, called restrict isometry property.