Basis Pursuit

I-Liang Chern

October 6, 2016

Basis Pursuit

- Given a vector $\mathbf{y}$ which is obtained from a sparse vector $\mathbf{x}$ through $\mathbf{y} = \mathbf{A}\mathbf{x}$. The original sparse recovery problem is to recover $\mathbf{x}$ by solving

$$\boxed{\text{(P0)} \quad \min_{\mathbf{z}} \|\mathbf{z}\|_1 \text{ subject to } \mathbf{A}\mathbf{z} = \mathbf{y}.}$$

- The basis pursuit is to solve

$$\boxed{\text{(P1)} \quad \min_{\mathbf{z}} \|\mathbf{z}\|_1 \text{ subject to } \mathbf{A}\mathbf{z} = \mathbf{y}.}$$

- Goal: To find equivalent condition on $\mathbf{A}$ so that solving (P1) is equivalent to solving (P0).[1]

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

Outline

- Null space property: recovery condition characterized by null space of $\mathbf{A}$
- Stability
- Robustness

Recovery condition in terms of $N(\mathbf{A})$

Definition

1. A matrix $\mathbf{A} \in \mathbb{C}^{m \times N}$ is said to satisfy the null space property relative to $S \subset [N]$ if

$$\|\mathbf{v}_S\|_1 < \|\mathbf{v}_{\bar{S}}\|_1 \text{ for all } \mathbf{v} \in N(\mathbf{A}) \setminus \{0\}.$$

2. It is said to satisfy the null space property of order s if it satisfies the null space property relative to S for all $S \subset [N]$ with $|S| \leq s$.

Example

1. $N = 2$, $m = 1$, $S = \{2\}$.

$$\mathbf{A} = (-1, 2), \quad -x_1 + 2x_2 = 2,$$

$$N(\mathbf{A}) =< \mathbf{v} >=< (2, 1) >, \quad |v_2| < |v_1|$$

2. $N = 3$, $m = 1$,

$$\mathbf{A} = \begin{pmatrix} 2 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}, \quad N(\mathbf{A}) =< (-1, 2, 2)^T > .$$

$$|v_1| < |v_2| + |v_3|, \quad |v_2| < |v_1| + |v_3|, \quad |v_3| < |v_1| + |v_2|$$

Null space property $\Leftrightarrow$ Exact recovery

Theorem

Given an $m \times N$ *matrix* $\mathbf{A}$, *every* N*-vector* $\mathbf{x}$ *supported on* S *is the unique solution of (P1) with* $\mathbf{y} = \mathbf{A}\mathbf{x}$ *if and only if* $\mathbf{A}$ *satisfies the null space property relative to* S.

Proof. NSP $\Rightarrow$ ExRy

1. Let $\mathbf{z}$ satisfy $\mathbf{A}\mathbf{z} = \mathbf{A}\mathbf{x}$. We want to show $\|\mathbf{x}\|_1 < \|\mathbf{z}\|_1$ if $\mathbf{A}$ satisfies NSP w.r.t. $S := \text{supp}\,(\mathbf{x})$.
2. Let $\mathbf{v} := \mathbf{x} - \mathbf{z}$. Then

$$\begin{aligned}\|\mathbf{x}\|_1 \leq \|\mathbf{x} - \mathbf{z}_S\|_1 + \|\mathbf{z}_S\|_1 &= \|\mathbf{v}_S\| + \|\mathbf{z}_S\|_1 \\ < \|\mathbf{v}_{\bar{S}}\|_1 + \|\mathbf{z}_S\|_1 &= \|\mathbf{z}_{\bar{S}}\|_1 + \|\mathbf{z}_S\|_1 = \|\mathbf{z}\|_1\end{aligned}$$

3 Uniqueness is followed by the strict inequality $\|\mathbf{x}\| < \|\mathbf{z}\|_1$ for all $\mathbf{Az} = \mathbf{Ax}$ and $\mathbf{z} \neq \mathbf{x}$.

Proof: ExRy $\Rightarrow$ NSP

1. For any $\mathbf{v} \in N(\mathbf{A}) - \{0\}$, $\mathbf{v}_S$ is the unique solution solving (P1) with $\mathbf{Az} = \mathbf{Av}_S$.
2. But we have $\mathbf{A}(-\mathbf{v}_{\bar{S}}) = \mathbf{Av}_S$. From ExRy, we get

$$\|\mathbf{v}_S\|_1 < \|\mathbf{v}_{\bar{S}}\|_1$$

Stability

- The data vector may not be sparse, it may have defect, or it may just be compressible.
- Compressibility of $\mathbf{x}$ is measured by $\sigma_s(\mathbf{x})_1 := \|\mathbf{x} - \mathbf{x}_s^*\|_1$, where $\mathbf{x}_s^*$ contains the largest s component (in magnitude) of $\mathbf{x}$.

Definition

A matrix $\mathbf{A}$ satisfied the stable null space property with constant $0 < \rho < 1$ relative to $S \subset [N]$ if

$$\|\mathbf{v}_S\|_1 \leq \rho \|\mathbf{v}_{\bar{S}}\|_1 \text{ for all } \mathbf{v} \in N(\mathbf{A}). \tag{0.1}$$

Stable CS

Theorem (Stable CS)

Suppose $\mathbf{A}$ *satisfies the stable null space property of order* s*. Then given any vector* $\mathbf{x}$*, a soultion* $\mathbf{x}^{\#}$ *of (P1) with* $\mathbf{y} = \mathbf{A}\mathbf{x}$ *satisfies*

$$\|\mathbf{x} - \mathbf{x}^{\#}\|_1 \leq \frac{2(1+\rho)}{1-\rho}\sigma_s(\mathbf{x})_1. \tag{0.2}$$

Remarks.

- If $\mathbf{x}$ is indeed s sparse, then $\mathbf{x}^{\#} = \mathbf{x}$.
- No uniqueness is required here.

Theorem

The matrix $\mathbf{A} \in \mathbb{C}^{m\times N}$ *satisfies the stable null space property:* $\exists\ 0 < \rho < 1$ *s.t.*

$$\|\mathbf{v}_S\|_1 \le \rho \|\mathbf{v}_{\bar{S}}\|_1 \quad \forall \mathbf{v} \in N(\mathbf{A}) \setminus \{0\},$$

if and only if

$$\|\mathbf{z} - \mathbf{x}\|_1 \le \frac{1+\rho}{1-\rho} \left(\|\mathbf{z}\|_1 - \|\mathbf{x}\|_1 + 2\|\mathbf{x}_{\bar{S}}\|_1 \right) \tag{0.3}$$

for any $\mathbf{z}$ *with* $\mathbf{Az} = \mathbf{Ax}$*.*

Proof of Stable CS Theorem.

In (0.3), take $\mathbf{z} = \mathbf{x}^{\#}$, then from $\mathbf{Ax}^{\#} = \mathbf{Ax}$ and $\|\mathbf{x}^{\#}\|_1 \le \|\mathbf{x}\|_1$, (0.2) follows.

Proof of Theorem 0.5

(0.3)⇒(0.1).

1. Given any $\mathbf{v} \in N(\mathbf{A}) \setminus \{0\}$, since $\mathbf{A}\mathbf{v}_{\bar{S}} = \mathbf{A}(-\mathbf{v}_S)$, we apply (0.3) with $\mathbf{x} = -\mathbf{v}_S$ and $\mathbf{z} = \mathbf{v}_{\bar{S}}$. It yields

$$\|\mathbf{v}\|_1 \le \frac{1+\rho}{1-\rho}\left(\|\mathbf{v}_{\bar{S}}\|_1 - \|\mathbf{v}_S\|_1\right).$$

2. This is the same as

$$(1-\rho)\left(\|\mathbf{v}_S\| + \|\mathbf{v}_{\bar{S}}\|_1\right) \le (1+\rho)\left(\|\mathbf{v}_{\bar{S}}\|_1 - \|\mathbf{v}_S\|_1\right)$$

which gives

$$\|\mathbf{v}_S\|_1 \le \rho\|\mathbf{v}_{\bar{S}}\|_1.$$

(0.1)⇒(0.3).

1. Suppose $\mathbf{z}$ satisfies $\mathbf{A}\mathbf{z} = \mathbf{A}\mathbf{x}$. Let $\mathbf{v} := \mathbf{z} - \mathbf{x}$.
2. From (0.1)

$$\|\mathbf{v}\|_1 = \|\mathbf{v}_S\|_1 + \|\mathbf{v}_{\bar{S}}\|_1 \leq (1+\rho)\|\mathbf{v}_{\bar{S}}\|_1.$$

3. We claim

$$\|\mathbf{v}_{\bar{S}}\|_1 \leq \|\mathbf{z}\|_1 - \|\mathbf{x}\|_1 + \|\mathbf{v}_S\|_1 + 2\|\mathbf{x}_{\bar{S}}\|_1$$

4. With this, use (0.1) again,

$$\|\mathbf{v}_{\bar{S}}\|_1 \leq \|\mathbf{z}\|_1 - \|\mathbf{x}\|_1 + \rho\|\mathbf{v}_{\bar{S}}\|_1 + 2\|\mathbf{x}_{\bar{S}}\|_1$$

$$\|\mathbf{v}_{\bar{S}}\|_1 \leq \frac{1}{1-\rho}\left(\|\mathbf{z}\|_1 - \|\mathbf{x}\|_1 + 2\|\mathbf{x}_{\bar{S}}\|_1\right)$$

5. (0.1) follows from 2 and 4.

Proof of claim:

1.

$$\|\mathbf{x}\|_1 = \|\mathbf{x}_{\bar{S}}\|_1 + \|\mathbf{x}_S\|_1 \leq \|\mathbf{x}_{\bar{S}}\|_1 + \|(\mathbf{x}-\mathbf{z})_S\|_1 + \|\mathbf{z}_S\|_1$$

$$\|(\mathbf{x}-\mathbf{z})_{\bar{S}}\|_1 \leq \|\mathbf{x}_{\bar{S}}\|_1 + \|\mathbf{z}_{\bar{S}}\|_1$$

2. Adding these two gives the claim

$$\|(\mathbf{x}-\mathbf{z})_{\bar{S}}\|_1 \leq \|\mathbf{z}\|_1 - \|\mathbf{x}\|_1 + \|(\mathbf{x}-\mathbf{z})_S\|_1 + 2\|\mathbf{x}_{\bar{S}}\|_1$$

Robustness

- Consider the case when there is measurement error:

$$\|\mathbf{y} - \mathbf{A}\mathbf{x}\| \leq \eta.$$

- We therefore consider thew minimization problem

$$\boxed{(P_{1\eta}) \quad \min \|\mathbf{z}\|_1 \text{ subject to } \|\mathbf{A}\mathbf{z} - \mathbf{y}\| \leq \eta}$$

Definition

$\mathbf{A}$ is said to satisfy the robust null space property with ρ and τ w.r.t. $S \subset [N]$ if

$$\|\mathbf{v}_S\|_1 \leq \rho \|\mathbf{v}_{\bar{S}}\|_1 + \tau \|\mathbf{A}\mathbf{v}\| \quad \forall \ \mathbf{v} \in \mathbb{C}^N.$$

Robustness CS

Theorem

Suppose $\mathbf{A}$ *satisfies robustness NSP for all* S *with* $|S| \leq s$*.*
Then for any $\mathbf{x}$*, if* $\mathbf{x}^{\#}$ *is a solution of* $(P_{1\eta})$ *with* $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{e}$ *and* $\|\mathbf{e}\| \leq \eta$*, then*

$$\|\mathbf{x}^{\#} - \mathbf{x}\|_1 \leq \frac{2(1+\rho)}{1-\rho}\sigma_s(\mathbf{x})_1 + \frac{4\tau}{1-\rho}\eta. \qquad (0.4)$$

Theorem

The matrix $\mathbf{A}$ *satisfies robustness NSP w.r.t* S

$$\|\mathbf{v}_S\|_1 \leq \rho\|\mathbf{v}_{\bar{S}}\|_1 + \tau\|\mathbf{A}\mathbf{v}\| \quad \forall\ \mathbf{v} \in \mathbb{C}^N. \tag{0.5}$$

if and only if

$$\|\mathbf{z}-\mathbf{x}\|_1 \leq \frac{1+\rho}{1-\rho}\left(\|\mathbf{z}\|_1 - \|\mathbf{x}\|_1 + 2\|\mathbf{x}_{\bar{S}}\|_1\right) + \frac{2\tau}{1-\rho}\|\mathbf{A}(\mathbf{z}-\mathbf{x})\| \tag{0.6}$$

for any vectors $\mathbf{x}, \mathbf{z} \in \mathbb{C}^N$*.*

1. (0.6)$\Rightarrow$ (0.5): For any $\mathbf{v} \in \mathbb{C}^N$, taking $\mathbf{x} = -\mathbf{v}_S$, $\mathbf{z} = \mathbf{v}_{\bar{S}}$,

$$\|\mathbf{v}\|_1 \leq \frac{1+\rho}{1-\rho}\left(\|\mathbf{v}_{\bar{S}}\|_1 - \|\mathbf{v}_S\|_1\right) + \frac{2\tau}{1-\rho}\|\mathbf{A}\mathbf{v}\|$$

   Rearranging this get (0.5).

2. (0.5)$\Rightarrow$ (0.6): Taking $\mathbf{v} = \mathbf{z} - \mathbf{x}$,

$$\|\mathbf{v}\|_1 = \|\mathbf{v}_S\|_1 + \|\mathbf{v}_{\bar{S}}\|_1 \leq (1+\rho)\|\mathbf{v}_{\bar{S}}\|_1 + \tau\|\mathbf{A}\mathbf{v}\|.$$

$$\begin{aligned}\|\mathbf{v}_{\bar{S}}\|_1 &\leq \|\mathbf{z}\|_1 - \|\mathbf{x}\|_1 + \|\mathbf{v}_S\|_1 + 2\|\mathbf{x}_{\bar{S}}\|_1 \\ &\leq \|\mathbf{z}\|_1 - \|\mathbf{x}\|_1 + \rho\|\mathbf{v}_{\bar{S}}\|_1 + \tau\|\mathbf{A}\mathbf{v}\| + 2\|\mathbf{x}_{\bar{S}}\|_1\end{aligned}$$

   Combine these two, we get (0.6).

Robust recovery in ℓ_2

Definition

$\mathbf{A}$ is said to satisfy the ℓ_2-robust null space property with $0 < \rho < 1$ and τ w.r.t. $S \subset [N]$ if

$$\|\mathbf{v}_S\|_2 \leq \frac{\rho}{\sqrt{s}}\|\mathbf{v}_{\bar{S}}\|_1 + \tau\|\mathbf{A}\mathbf{v}\|_2 \quad \forall\ \mathbf{v} \in \mathbb{C}^N.$$

Remark. Please compare this with the robust NSP defined earlier:

$$\|\mathbf{v}_S\|_1 \leq \rho\|\mathbf{v}_{\bar{S}}\|_1 + \tau\|\mathbf{A}\mathbf{v}\|_2 \quad \forall\ \mathbf{v} \in \mathbb{C}^N.$$

Also, we have Hölder inequality $\|\mathbf{v}_S\|_1 \leq \sqrt{s}\|\mathbf{v}_S\|_2$.

Theorem

Suppose $\mathbf{A}$ satisfies ℓ_2-robust NSP with $0 < \rho < 1$ and τ for all S with $|S| \le s$. Then for any $\mathbf{x}$, if $\mathbf{x}^{\#}$ is a solution of $(P_{1\eta})$ with $\mathbf{y} = \mathbf{Ax} + \mathbf{e}$ and $\|\mathbf{e}\| \le \eta$, then

$$\|\mathbf{x}^{\#} - \mathbf{x}\|_1 \le C\sigma_s(\mathbf{x})_1 + D\sqrt{s}\eta. \tag{0.7}$$

$$\|\mathbf{x}^{\#} - \mathbf{x}\|_2 \le \frac{C}{\sqrt{s}}\sigma_s(\mathbf{x})_1 + D\eta. \tag{0.8}$$

where C, D depend on ρ and τ.

Theorem

The matrix $\mathbf{A}$ *satisfies* ℓ_2*-robust NSP w.r.t* S

$$\|\mathbf{v}_S\|_2 \leq \frac{\rho}{\sqrt{s}}\|\mathbf{v}_{\bar{S}}\|_1 + \tau\|\mathbf{A}\mathbf{v}\|_2 \quad \forall\ \mathbf{v} \in \mathbb{C}^N. \tag{0.9}$$

then for $1 \leq p \leq 2$*,*

$$\|\mathbf{z}-\mathbf{x}\|_p \leq \frac{C}{s^{1-1/p}}\left(\|\mathbf{z}\|_1 - \|\mathbf{x}\|_1 + 2\|\mathbf{x}_{\bar{S}}\|_1\right) + Ds^{1/p-1/2}\|\mathbf{A}(\mathbf{z}-\mathbf{x})\|_2 \tag{0.10}$$

where

$$C = \frac{(1+\rho)^2}{1-\rho}, \quad D = \frac{(3+\rho)\tau}{1-\rho}.$$

Proof.

1. By Hölder inequality
$$\|\mathbf{v}_S\|_p \le s^{1/p-1/q}\|\mathbf{v}_S\|_q$$
In particular, $q = 2$, together with ℓ_2-robust NSP, we get
$$\|\mathbf{v}_S\|_p \le s^{1/p-1/2}\|\mathbf{v}_S\|_2 \le s^{1/p-1}\rho\|\mathbf{v}_{\bar{S}}\|_1 + s^{1/p-1/2}\tau\|\mathbf{A}\mathbf{v}\|_2.$$
2. With $p = 1$, apply (0.6), we get
$$\|\mathbf{z}-\mathbf{x}\|_1 \le \frac{1+\rho}{1-\rho}\left(\|\mathbf{z}\|_1 - \|\mathbf{x}\|_1 + 2\|\mathbf{x}_{\bar{S}}\|_1\right) + \frac{2\tau}{1-\rho}s^{1/2}\|\mathbf{A}(\mathbf{z}-\mathbf{x})\|$$
3. Take S the index set of s largest entries of $\mathbf{v} = \mathbf{z} - \mathbf{x}$,
$$\begin{aligned}\|\mathbf{z}-\mathbf{x}\|_p &\le \|(\mathbf{z}-\mathbf{x})_{\bar{S}}\|_p + \|(\mathbf{z}-\mathbf{x})_S\|_p \le \frac{1}{s^{1-1/p}}\|\mathbf{z}-\mathbf{x}\|_1 + \|(\mathbf{z}-\mathbf{x})_S\|_p \\ &\le \frac{1}{s^{1-1/p}}\|\mathbf{z}-\mathbf{x}\|_1 + s^{1/p-1}\rho\|(\mathbf{z}-\mathbf{x})_{\bar{S}}\|_1 + s^{1/p-1/2}\tau\|\mathbf{A}(\mathbf{z}-\mathbf{x})\|_2 \\ &\le \frac{1+\rho}{s^{1-1/p}}\|\mathbf{z}-\mathbf{x}\|_1 + s^{1/p-1/2}\tau\|\mathbf{A}(\mathbf{z}-\mathbf{x})\|_2\end{aligned}$$

Recall $\boxed{\sigma_s(\mathbf{x})_q \le \frac{c_{p,q}}{s^{1/p-1/q}}\|\mathbf{x}\|_p,}$