

Chapter 8

Sparse Recovery

- High Dimensional Signal Recovery
- signal recovery based on M^* bound
- exact recovery based on escape
- deterministic conditions for recovery
- Recovery with noise (LASSO/BPDN)

Problem: Fix a vector $x_{\#} \in \mathbb{R}^n$ and we get to see $y \in \mathbb{R}^m$ satisfying

$$y = Ax_{\#} + w$$

where $w \in \mathbb{R}^m$ is an error vector.

Notation

- $x_{\#}$ is called the signal
- y is called the measurement
- A is called the design matrix
- w is called the noise

The diagram shows the equation $y = Ax_{\#} + w$ using matrix notation. On the left, a vertical rectangle represents the measurement vector y , with the label m to its left. This is followed by an equals sign. In the center, a horizontal rectangle represents the design matrix A , with the label n above it and A inside. To the right of A is a vertical rectangle representing the signal vector $x_{\#}$, with the label $x_{\#}$ to its left. This is followed by a plus sign. On the far right, another vertical rectangle represents the noise vector w , with the label w inside and m to its right.

We'll be interested in setting
 under additional prior information $x_{\#} \in T$
 $m \ll n$
 For the time being assume $w=0$.

Summary:

Having access to y , find $x \in \mathbb{R}^n$ s.t.
 (*) $y = Ax$ and $x \in T$.

Thm: Suppose $A \in \mathbb{R}^{m \times n}$ has independent,
 isotropic, σ -subGaussian rows. Then any
 point \hat{x} satisfying (*) satisfies

$$\mathbb{E} \|\hat{x} - x_{\#}\|_2 \leq \frac{C\sigma^2\omega(T)}{\sqrt{m}}$$

pf: Since $\hat{x}, x_{\#} \in T$ satisfy $A\hat{x} = Ax_{\#} = y$,
 we know $\hat{x} \in x_{\#} + \ker A$. So m^* bound

$$\begin{aligned} \Rightarrow \mathbb{E} \|\hat{x} - x_{\#}\|_2 &\leq \mathbb{E} \text{diam}((x_{\#} + \ker(A)) \cap T) \\ &\leq \frac{C\sigma^2\omega(T)}{\sqrt{m}} \quad \square \end{aligned}$$

Remark: So can achieve

$$\mathbb{E} \|\hat{x} - x_{\#}\|_2 \leq \varepsilon \text{diam}(T)$$

using $m \geq C \frac{\sigma^4}{\varepsilon^2} d(T)$ measurements.

Often prior information encodes sparsity.

Define $\|x\|_0 = |\{i: x_i \neq 0\}|$

Suppose it's the case that

$$\|x_{\#}\|_0 =: s \ll n$$

This corresponds to setting

$$T = \{x \in \mathbb{R}^n: \|x\|_0 \leq s\}$$

Finding a point x satisfying

$$Ax = y \text{ and } x \in T \quad \textcircled{1}$$

can be difficult.

① It can be done efficiently under RTP assumptions by alternating projections

Instead, let's use a convex surrogate
Replace $\|x\|_0$ by $\|x\|_1$.

Intuition

Lim $\|x\|_p = \|x\|_0$ and $\|x\|_1$ is the first true norm.

How big should the l_1 -ball be?

Lemma: $\{x \in \mathbb{R}^n : \|x\|_0 \leq s, \|x\|_2 \leq 1\} \subseteq \sqrt{s} B_1^n$

pt: $\sum_{i=1}^n |x_i| \leq \sum_{i \in \text{supp } x} |x_i| \leq \sqrt{s} \|x\|_2 \leq \sqrt{s}$ \square

So new plan

① Find x : $y = Ax$ and $x \in \sqrt{s} B_1^n$

Thm: Assume $\|x_\# \|_0 \leq s$ and $\|x_\# \|_2 \leq 1$

Then any solution of ① satisfies

$$\mathbb{E} \|\hat{x} - x_\# \|_2 \leq c \sigma^2 \sqrt{\frac{s \log(n)}{m}}$$

pt: Set $T = \sqrt{s} B_1^n$. Then $x_\#, \hat{x} \in T$ and therefore

$$\mathbb{E} \|\hat{x} - x_\# \|_2 \leq \frac{c \sigma^2 \omega(T)}{\sqrt{m}} = c \sigma^2 \sqrt{\frac{s \log n}{m}} \quad \square$$

So recovery is possible if
 $m \sim s \log(n)$

Remark: dependence can be improved to
 $m \sim s \log(n/s)$

Notice that $\textcircled{1}$ requires to know s and the theorem requires $\|x_{\#}\|_2 \leq 1$. We can do better.

Thm: Suppose $\|x_{\#}\|_0 \leq s$. Then any solution \hat{x} of

$$\begin{array}{l} \min_x \|x\|_1 \\ \text{s.t. } Ax = y \end{array}$$

satisfies

$$\mathbb{E} \|\hat{x} - x_{\#}\|_2 \leq c \sqrt{\frac{s \log n}{m}} \|x_{\#}\|_2$$

pf: We know $A\hat{x} = Ax = y$

$$\text{Also } \|x_{\#}\|_1 \leq \sqrt{s} \cdot \|x_{\#}\|_2$$

$$\Rightarrow \|\hat{x}\|_1 \leq \sqrt{s} \cdot \|x_{\#}\|_2$$

$\hat{x}, x_{\#}$ satisfy $Ax' = y$, $x' \in \sqrt{s} \|x_{\#}\|_2 B_1^n$

$$\Rightarrow \mathbb{E} \|\hat{x} - x_{\#}\|_2 \leq c \sigma^2 \sqrt{\frac{s \log n}{m}} \cdot \|x_{\#}\|_2$$

The previous results extend easily to low-rank matrix recovery.

Model: There is $X \in \mathbb{R}^{d \times d}$ satisfying

$$y_i = \langle A_i, X \rangle \quad \text{for } i=1, \dots, m$$

where $A_i \in \mathbb{R}^{d \times d}$ are independent.

The side information we assume is $\text{rank}(X) =: r \ll d$

Let $s(X)$ be the singular values of X .

Then $\text{rank}(X) = \|s(X)\|_0$

$$\sum_{i=1}^d s_i(X) = \|X\|_* = \|s(X)\|_1$$

Lemma: $\{X : \text{rank}(X) \leq r, \|X\|_F \leq 1\} \subset \sqrt{r} B_*$

pt: $\|X\|_* = \sum_{i=1}^r s_i(X) \leq \sqrt{r} \|s(X)\|_2 = \sqrt{r} \|X\|_F \leq \sqrt{r} \quad \square$

Lemma: $\omega(B_{\infty}^d) \leq c\sqrt{d}$

pf: $\mathbb{E} \sup_{\|X\|_{\infty} \leq 1} \langle G, X \rangle = \mathbb{E} \|G\|_2 \leq c\sqrt{d}$ \square

Thm: Suppose that A_i are independent and $\text{vec}(A_i)$ are σ -subgaussian, ~~isotype~~
Assume $X_{\#}$ satisfies $\|X_{\#}\|_F \leq 1$.

Then any \hat{X} satisfying

$y_i = \langle A_i, \hat{X} \rangle \quad \forall i=1, \dots, m, \quad \|\hat{X}\|_{\infty} \leq \sqrt{r}$
satisfies

$$\mathbb{E} \|\hat{X} - X_{\#}\|_F \leq c\sigma^2 \sqrt{\frac{rd}{m}}$$

pf: Set $T = \sqrt{r} B_{\infty}^d$. Then we can treat X, \hat{X} as long vectors and then

$$\mathbb{E} \|\hat{X} - X_{\#}\|_F \leq \frac{C\sigma^2 \omega(T)}{\sqrt{m}} = C\sigma^2 \sqrt{\frac{rd}{m}} \quad \square$$

As in the vector case, we can instead solve

$$\begin{array}{ll} \min & \|X\|_* \\ \text{s.t.} & y_i = \langle A_i, X \rangle \quad \forall i=1, \dots, m. \end{array}$$

Exact Recovery Guarantees:

Consider the sparse recovery problem

$$\begin{array}{ll} \min & \|x\|_1 \\ \text{s.t.} & Ax = b \end{array}$$

We will now show that with h.p. $\hat{x} = x_{\#}$.

Convex Optimization:

$$\text{Consider } \begin{array}{ll} \min & f(x) \\ \text{s.t.} & Ax = b \end{array}$$

where f is convex. Then \bar{x} is optimal

$$\text{iff } \left\{ \begin{array}{l} A\bar{x} = b \\ f'(\bar{x}; v) \geq 0 \quad \forall v \in \ker A \end{array} \right\}$$

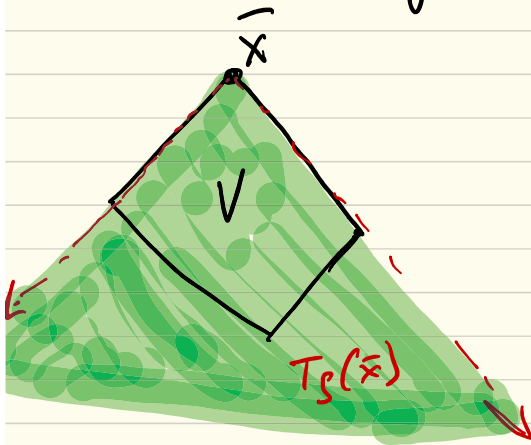
For any convex set $V \subseteq \mathbb{R}^n$ and $\bar{x} \in V$ define the tangent cone $T_V(\bar{x}) = \text{cl} \bigcup_{t>0} t(V - \bar{x})$

Lemma:

As long as \bar{x} is not a minimizer of f ,
have

$$T_V(\bar{x}) = \{v: f'(\bar{x}, v) \leq 0\}$$

$$\text{where } V = \{y: f(y) \leq f(\bar{x})\}$$



So conditions

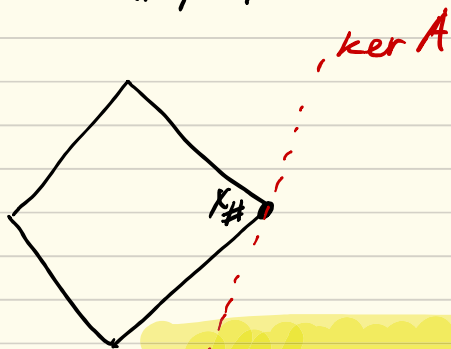
$$\left\{ \begin{array}{l} A\bar{x} = b \\ T_{[f(x) \leq f(\bar{x})]}(\bar{x}) \cap \ker(A) = \{0\} \end{array} \right\}$$

guarantees \bar{x} is a soln, and is in fact the unique soln.

So to show that $x_{\#}$ is the unique solution of

$\min \|x\|,$
s.t. $Ax=b$
we estimate probability

$$T_{\|x_{\#}\|, B, (x_{\#})} \cap \ker A = \{0\}$$



Define $Q := T_{\|x_{\#}\|, B, (x_{\#})}$

Recall by the escape thm, the probability is governed by

$$\omega(Q \cap S^{n-1})$$

Lemma: $\|h\|_1 \leq 2\sqrt{s} \|h\|_2 \quad \forall h \in Q$.

pf: Let $S = \text{supp}(x_{\#})$. Let's verify

$$\|h_S\|_1 \leq \|h_S\|_1, \quad \forall h \in Q$$

Let $x \in \|x_{\#}\|_1 B_1$, and define $h = x - x_{\#}$.

Then

$$\begin{aligned} \|x_{\#}\|_1 &\geq \|x\|_1 = \|x_{\#} + h\|_1 = \|(x_{\#})_S + h_S\|_1 + \|(x_{\#})_{S^c} + h_{S^c}\|_1 \\ &\geq \|(x_{\#})_S\|_1 - \|h_S\|_1 + \|h_{S^c}\|_1 \end{aligned}$$

$$\Rightarrow \|h_S\|_1 \geq \|h_{S^c}\|_1 \quad \text{ok!}$$

$$\text{So } \|h\|_1 = \|h_S\|_1 + \|h_{S^c}\|_1 \leq 2\|h_S\|_1 \leq 2\sqrt{s} \|h\|_2 \quad \square$$

So we learned

$$Q \cap B_2 \subseteq 2\sqrt{s} B_1$$

Thm: Suppose rows of $A \in \mathbb{R}^{m \times n}$ are independent, isotropic, σ -subGaussian. Then as long as

$$m \geq c\sigma^4 s \log n$$

then with probability $1 - 2\exp(-\frac{cm}{\sigma^4})$, $x_{\#}$ is the unique soln of

$$\begin{aligned} \min_x \|x\|, \\ \text{s.t. } Ax = b \end{aligned}$$

pf: The escape thm shows that as long as

$$m \geq c\sigma^4 \omega(Q \cap S^{n-1})$$

then

$$(Q \cap S^{n-1}) \cap \ker A = \emptyset$$

w.p. $1 - 2\exp(-cm/\sigma^4)$.

From the lemma

$$\omega(Q \cap B_2) \leq c\omega(\sqrt{s} B_1) \leq c\sqrt{s \log(n)}$$

We next look at deterministic conditions for exact recovery.

Looking back at the previous proof, define

$$\mathcal{C}(S) := \{h \in \mathbb{R}^n : \|h_{S^c}\|_1 \leq \|h_S\|_1\}$$

Recall we showed:

$$T_{\|x_{\#}\|_1, B_1}(x_{\#}) \subseteq \mathcal{C}(S)$$

where $S = \text{supp}(x_{\#})$.

Defn: $A \in \mathbb{R}^{m \times n}$ satisfies the restricted nullspace property with respect to

$S \subseteq \{1, \dots, n\}$ if $\mathcal{C}(S) \cap \ker(A) = \{0\}$.

Thm: The following properties are equivalent.
① $A \in \mathbb{R}^{m \times n}$ satisfies the restricted nullspace property with respect to S

② for any $x_{\#} \in \mathbb{R}^n$ with $\text{supp}(x_{\#}) = S$,
the problem

$$\begin{array}{ll} \min & \|x\|_1 \\ \text{s.t.} & Ax = b \end{array}$$

has the unique soln $x_{\#}$

pf: We already proved ① \Rightarrow ②

To see converse, fix $\bar{x} \in \ker A \setminus \{0\}$.

Then

$$\begin{array}{ll} \min & \|x\|_1 \\ \text{s.t.} & Ax = A \begin{bmatrix} \bar{x}_S \\ 0 \end{bmatrix} \end{array}$$

has $\begin{bmatrix} \bar{x}_S \\ 0 \end{bmatrix}$ as the unique optimal soln.

Since $A\bar{x} = 0$, it follows $\begin{bmatrix} 0 \\ -\bar{x}_{S^c} \end{bmatrix}$ is also feasible.

$\Rightarrow \|\bar{x}_S\|_1 < \|\bar{x}_{S^c}\|_1 \Rightarrow \bar{x} \notin \mathcal{L}(S) \quad \square$

Recall[‡] we also showed

$$\|h\|_1 \leq 2\sqrt{S} \|h\|_2 \quad \forall h \in \mathcal{C}(S)$$

So $\mathcal{C}(S) \cap B_2 \subset 2\sqrt{S} B_1$,

and escape then show that if $A \in \mathbb{R}^{m \times n}$ is random (independent, subGaussian, isotropic rows) and

$$m \geq c \sigma^4 \sqrt{S} \log(n)$$

then w.p. $1 - 2 \exp(-\frac{cm}{\sigma^4})$ have

$$\mathcal{C}(S) \cap \ker A = \{0\}$$

Recall $\forall h \in \mathcal{C}(S)$

$$\|h\|_1 = \|h_S\|_1 + \|h_{S^c}\|_1 \leq 2\|h\|_S \leq 2\sqrt{|S|} \|h\|_2 \quad \ddagger$$

Q: How to verify the restricted nullspace property?

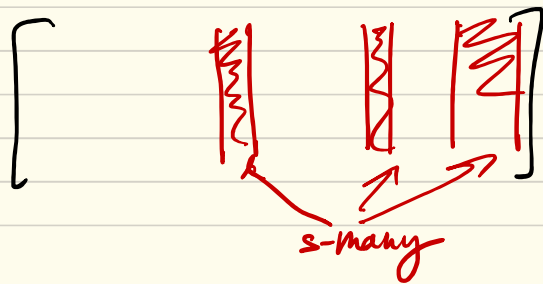
Here's a simpler condition

Defn: $A \in \mathbb{R}^{m \times n}$ satisfies the restricted isometry property relative to $S \subseteq \{1, \dots, n\}$ with constant $\alpha, \beta > 0$ if

$$\alpha \|x\|_2 \leq \frac{1}{\sqrt{|S|}} \|Ax\|_2 \leq \beta \|x\|_2$$

for all x with $\|x\|_0 \leq s$.

Exercise: This is equivalent to requiring $\{\text{singular values of } \frac{1}{\sqrt{|S|}} A_S\} \subset [\alpha, \beta]$ where A_S is submatrix of columns S and $|S| \leq s$.



Thm.: If $\frac{2}{3} < \alpha < \beta < \frac{4}{3}$, then A satisfies the restricted nullspace property for any $S \subseteq \{1, \dots, n\}$ with $|S| \leq s$
[See Wainwright Proposition 7.11]

Thm.: Suppose $A \in \mathbb{R}^{m \times n}$ satisfies RIP with parameters $\alpha, \beta, (1+\lambda)s$ where $\lambda > (\frac{\beta}{\alpha})^2$. Then every s -sparse $x \in \mathbb{R}^n$ can be recovered exactly by solving $\min \|x\|$,
s.t. $Ax = b$ (*)

pf.: Define $h = \hat{x} - x$ where \hat{x} solves (*).
Let $S_0 = \text{supp}(x)$, let S_1 index λs largest elements of $h_{S_0^c}$ in absolute value, let S_2 index the next λs largest in $h_{S_0^c}$, and so on.

Let $S_{0,1} = S_0 \cup S_1$.

Since $Ah = Ax - Ax_0 = 0$, we have

$$0 = \|Ah\|_2 \geq \|A_{S_{0,1}} h_{S_{0,1}}\|_2 - \|A_{S_{0,1}^c} h_{S_{0,1}^c}\|_2$$

Since $|S_{0,1}| \leq s + \lambda s$, RIP gives

$$\|A_{S_{0,1}} h_{S_{0,1}}\|_2 \geq \alpha \sqrt{m} \cdot \|h_{S_{0,1}}\|_2$$

and

$$\|A_{S_{0,1}^c} h_{S_{0,1}^c}\|_2 \leq \sum_{i \geq 2} \|A_{S_i} h_{S_i}\|_2$$

$$\leq \beta \sqrt{m} \sum_{i \geq 2} \|h_{S_i}\|_2$$

S_0

$$\alpha \|h_{S_{0,1}}\|_2 \leq \beta \sum_{i \geq 2} \|h_{S_i}\|_2$$

Notice

$$\{ \text{each entry of } h_{S_j} \} \leq \frac{\|h_{S_{j-1}}\|}{\lambda S}$$

for $j \geq 2$. [For simplicity assume]
n divides λS]

$$\text{So } \|h_{S_j}\|_2 \leq \frac{1}{\sqrt{\lambda S}} \|h_{S_{j-1}}\|_1$$

$$\Rightarrow \sum_{i \geq 2} \|h_{S_i}\|_2 \leq \frac{1}{\sqrt{\lambda S}} \sum_{i \geq 1} \|h_i\|_1$$

$$= \frac{1}{\sqrt{\lambda S}} \|h_{S_0}\|_1$$

$$\leq \frac{1}{\sqrt{\lambda S}} \|h_{S_0}\|_1 \quad [\mathcal{C}(S) \text{ contains}]$$

$$\leq \frac{1}{\sqrt{\lambda}} \|h_{S_0}\|_2$$

$$\leq \frac{1}{\sigma \lambda} \|h_{S_{0,1}}\|_2$$

$$\Rightarrow \frac{\beta}{\sqrt{\lambda}} \|h_{S_{0,1}}\|_2 \geq \alpha \|h_{S_{0,1}}\|_2 \Rightarrow h=0$$

Thm: Let $A \in \mathbb{R}^{m \times n}$ have independent, isotropic, σ -subgaussian rows. Assume

$$m \geq \frac{\sigma^4 s}{\delta^2} \log(en/s)$$

Then with probability $1 - 2 \exp(-\frac{c \delta^2 m}{\sigma^4})$ have

$$(1-\delta) \|x\|_2 \leq \frac{1}{\sqrt{m}} \|Ax\|_2 \leq (1+\delta) \|x\|_2 \quad \forall x \text{ with } \|x\|_2 \leq s.$$

pf: We will control singular values of A_S uniformly over all S' with $|S'|=s$.

For any fixed S , we know

$$\sqrt{m}-t \leq \sigma_n(A_S) \leq \sigma_1(A_S) \leq \sqrt{m}+t$$

with probability $1 - 2 \exp(-t^2)$ where $t = c\sigma^2(\sqrt{s}+t)$. Ensure

$$\frac{t}{\sqrt{m}} \leq \delta \text{ by setting } t = \frac{\delta \sqrt{m}}{c\sigma^2} - \sqrt{s}$$

Then $1-\delta \leq \sigma_n(A_{S'}) \leq \sigma_1(A_{S'}) \leq 1+\delta \quad \forall S': |S'|=s$

w.p.

$$1 - 2 \binom{n}{s} \exp(-t^2) \leq 1 - 2 \exp(s \log(\frac{en}{s}) - t^2) \leq 1 - 2 \exp(-\frac{c \delta^2 m}{64}) \quad \square$$

Sparse Recovery in the noisy setting

Suppose now that $x_{\#}$ satisfies

$$y = Ax_{\#} + w \quad \text{and} \quad \|x_{\#}\|_0 \leq s$$

and the noise vector $w \in \mathbb{R}^m$ is nonzero.

Three Problem Formulations: $S = \text{supp}(x_{\#})$

Ⓐ LASSO ("Least Absolute Shrinkage and Selection Operator")

$$\hat{x}_a \in \arg \min_x \left\{ \frac{1}{2m} \|y - Ax\|_2^2 + \lambda_m \|x\|_1 \right\}$$

mistfit model complexity

Ⓑ Basis Pursuit DeNoise (BPDN I)

$$\hat{x}_b \in \arg \min_x \frac{1}{2m} \|y - Ax\|_2^2 \quad \text{subject to} \quad \|x\|_1 \leq R$$

Ⓒ Basis Pursuit DeNoise (BPDN II)

$$\hat{x}_c \in \arg \min_x \|x\|_1 \quad \text{subject to} \quad \frac{1}{2m} \|y - Ax\|_2^2 \leq b^2$$

Remark: Solutions are the same for appropriate choices of (λ_m, R, b) but one can not easily set R from b etc.

We will need to assume a modified restricted nullspace property. Define

$$\mathcal{C}_\alpha(S) := \{h \in \mathbb{R}^d : \|h_S\|_1 \leq \alpha \|h_{S^c}\|_1\}$$

for some $\alpha \geq 0$. Note $T_{\|x_{\#}\|_1, B_1(x_{\#})} \subset \mathcal{C}_1(\text{supp}(x_{\#}))$

Defn: $A \in \mathbb{R}^{m \times n}$ satisfies the restricted eigenvalue (RE) condition with parameters (α, α) if

$$\frac{1}{m} \|Ah\|_2^2 \geq \alpha \|h\|_2^2 \quad \forall h \in \mathcal{C}_\alpha(S)$$

Remark:

$(\alpha, 1)$ is the restricted nullspace property.

Intuition: $\nabla^2 \left(\frac{1}{2m} \|b - Ax\|_2^2 \right) = \frac{1}{m} A^T A$ singular

RE ensures ∇^2 is positive definite in important directions.

Thm: Suppose $A \in \mathbb{R}^{m \times n}$ has independent, isotropic, σ -subGaussian rows. Then

$$\frac{1}{\sqrt{m}} \|Ah\|_2 \geq \left(1 - \delta - C\sigma^2(1+\alpha) \sqrt{\frac{s \log n}{m}}\right) \|h\|_2 \quad \forall h \in \mathcal{C}_s(\mathcal{S})$$

w.p. $1 - 2 \exp\left(-\frac{C\delta^2 m}{\sigma^4}\right)$

pf: Matrix Deviation with $T = \mathcal{C}_s(\mathcal{S}) \cap \mathcal{S}^{n-1}$ is

$$\sup_{x \in T} \left| \|Ax\|_2 - \sqrt{m} \|x\|_2 \right| \leq C\sigma^2(\omega(T) + u) \quad (*)$$

w.p. $1 - 2 \exp(-u^2)$.

Observe for $h \in T$ have

$$\|h\|_1 = \|h_S\|_1 + \|h_{S^c}\|_1 \leq (1+\alpha) \|h_S\|_1 \leq (1+\alpha) \sqrt{s} \cdot \|h_S\|_2 \leq 1$$

$$\Rightarrow \omega(T) \leq C(1+\alpha) \sqrt{s \log(n)}$$

$$\text{So } \frac{1}{\sqrt{m}} \|Ax\|_2 \geq \left(1 - C\sigma^2 \left((1+\alpha) \sqrt{\frac{s \log n}{m}} + \frac{u}{\sqrt{m}} \right)\right)$$

w.p. $1 - 2 \exp(-u^2)$. Set $u = \frac{\delta \sqrt{m}}{C\sigma^2}$ \square

Thm: Suppose $\text{supp}(x_{\#}) = S'$ with $\|x_{\#}\|_0 \leq S$, and A satisfies the RE condition (2.3)

(a) If $\lambda_m \geq 2 \left\| \frac{A^T w}{m} \right\|_{\infty}$, then

$$\|\hat{x}_a - x_{\#}\|_2 \leq \frac{3}{\sqrt{\kappa}} \sqrt{S} \lambda_m \stackrel{\text{if (a)}}{\approx} \frac{6}{\sqrt{\kappa}} \sqrt{S} \left\| \frac{A^T w}{m} \right\|_{\infty}$$

(b) If $R = \|x_{\#}\|_1$, then

$$\|\hat{x}_b - x_{\#}\|_2 \leq \frac{4}{\sqrt{\kappa}} \sqrt{S} \left\| \frac{A^T w}{m} \right\|_{\infty}$$

(c) If $b^2 \geq \frac{\|w\|_2^2}{2m}$, then

$$\|\hat{x}_c - x_{\#}\|_2 \leq \frac{4}{\sqrt{\kappa}} \sqrt{S} \left\| \frac{A^T w}{m} \right\|_{\infty} + \frac{2}{\sqrt{\kappa}} \sqrt{b^2 - \frac{\|w\|_2^2}{2m}}$$

In all cases $\|\hat{x} - x_{\#}\|_1 \leq 4\sqrt{S} \|\hat{x} - x_{\#}\|_2$

Example: Suppose $w \sim N(0, \sigma^2)$ and A is Gaussian

(a) Set $\lambda_m \approx \sigma \left(\sqrt{\frac{\log n}{m}} + \delta \right)$

$$\Rightarrow \|\hat{x}_1 - x_{\#}\|_2 \approx \frac{\sigma}{\sqrt{\kappa}} \sqrt{S} \left(\sqrt{\frac{\log n}{m}} + \delta \right)$$

w.p. $1 - 2 \exp\left(-\frac{m\delta^2}{2}\right)$

(b), (c) are similar. λ and b are easy to estimate.

Example: Suppose $A_{ij} \sim N(0,1)$ and $\|w\|_\infty \leq 5$

Same guarantees for LASSO:

pt: (b) Let $R = \|x_\# \|_1$, and let \hat{x} be a minimizer. Then

$$\frac{1}{2m} \|y - A\hat{x}\|_2^2 \leq \frac{1}{2m} \|y - Ax_\#\|_2^2$$

Plugging in $y = Ax_\# + w$

$$\Rightarrow \frac{\|Ah\|_2^2}{m} \leq \frac{2w^T Ah}{m} \quad \text{where } h = x - x_\#.$$

Since $h \in T_{RB}(x_\#)$ we know $h \in \mathcal{C}_1(S) \subseteq \mathcal{C}_2(S)$

$$\Rightarrow \|h\|_2^2 \leq 2\|h\|_1 \cdot \left\| \frac{A^T w}{m} \right\|_\infty \leq 2\sqrt{5} \|h\|_2 \left\| \frac{A^T w}{m} \right\|_\infty \quad \square$$

(c) Observe $\frac{1}{2m} \|y - Ax_\#\|_2^2 = \frac{\|w\|_2^2}{2m} \leq b^2$

So $x_\#$ is feasible $\Rightarrow \|\hat{x}\|_1 \leq \|x_\#\|_1$,

$\Rightarrow h = \hat{x} - x_\# \in \mathcal{C}_1(S)$.

Observe $\frac{1}{2m} \|y - A\hat{x}\|_2^2 \leq b^2 = \frac{1}{2m} \|y - Ax_\#\|_2^2 + \left(b - \frac{\|w\|_2}{m} \right)^2$

Rearranging

$$\frac{\|Ah\|_2^2}{m} \leq 2 \frac{w^T Ah}{m} + 2 \left(b^2 - \frac{\|w\|_2^2}{2m} \right)$$

$$\Rightarrow \kappa \|h\|_2^2 \leq 4\sqrt{5} \|h\|_2 \left\| \frac{Aw}{m} \right\|_\infty + 2 \left(b^2 - \frac{\|w\|_2^2}{2m} \right)$$

$$\leq \max \left\{ 4\sqrt{5} \|h\|_2 \left\| \frac{Aw}{m} \right\|_\infty, 2 \left(b^2 - \frac{\|w\|_2^2}{2m} \right) \right\}$$

(a) Define

$$L(x, \lambda_m) = \frac{1}{2m} \|y - Ax\|_2^2 + \lambda_m \|x\|_1$$

We know

$$L(\hat{x}, \lambda_m) \leq L(x_{\#}, \lambda_m) = \frac{1}{2} \|w\|_2^2 + \lambda_m \|x_{\#}\|_1$$

$$\Rightarrow \frac{1}{2m} \|Ah\|_2^2 \leq \frac{w^T Ah}{m} + \lambda_m (\|x_{\#}\|_1 - \|\hat{x}\|_1) \quad (*)$$

where $h = \hat{x} - x_{\#}$. We can write

$$\|x_{\#}\|_1 - \|\hat{x}\|_1 = \|x_{\#s}\|_1 - \|x_{\#s} + h_s\|_1 - \|h_{sc}\|_1$$

$$\leq \|h_s\|_1 - \|h_{sc}\|_1$$

$$\text{So } 0 \leq \frac{1}{m} \|Ah\|_2^2 \leq 2\|h\|_1 \left\| \frac{Aw}{m} \right\|_\infty + 2\lambda_m (\|h_s\|_1 - \|h_{sc}\|_1)$$

$$\text{choice of } \lambda_m \leq \lambda_m (\|h\|_1 + 2\|h_s\|_1 - 2\|h_{sc}\|_1) \leq \lambda_m (3\|h_s\|_1 - \|h_{sc}\|_1)$$

So $h \in \mathcal{C}_3(S)$. RE condition gives

$$\kappa \|h\|_2^2 \leq 3 \lambda_m \sqrt{S} \|h\|_2 \quad \square$$

In fact we proved bounds on MSE. Namely for (a):

$$\frac{1}{2m} \|A(\hat{x} - x_{\#})\|_2^2 \leq \frac{c \sqrt{S}}{\kappa} \|h\|_2 \left\| \frac{A^T w}{m} \right\|_{\infty}$$

$$\leq \frac{c S}{\kappa^2} \left\| \frac{A^T w}{m} \right\|_{\infty}^2$$

So if $w \sim N(0, I)$, get

$$\preceq \frac{c S \log(d)}{\kappa^2 m}$$

w.h.p.

$$\frac{1}{2m} \|y - A\hat{x}\|_2^2 + \lambda_m \|\hat{x}\|_1 \leq \frac{1}{2m} \|w\|_2^2 + \lambda_m \|x_{\#}\|_1$$

$\|Ax_{\#} + w\|_2^2$

$$\frac{1}{2m} \|w - A(\hat{x} - x_{\#})\|_2^2 + \lambda_m \|\hat{x}\|_1$$

$\|h\|_2^2$

$$\frac{1}{2m} \|w\|_2^2 - \frac{w^T Ah}{m} + \frac{1}{2m} \|Ah\|_2^2 + \lambda_m \|\hat{x}\|_1$$

isolate on left