04/26

$$\left[\begin{array}{l} \min f_0(x) : Ax = y = Ax_0 \\ x_0 \text{ is a minimizer} \Leftarrow \exists \lambda : \eta = A^T\lambda \in \partial f_0(x_0) \end{array}\right.$$

Can we certify that x_0 is the unique minimizer? Ref. Foucart, Rauhut book.

ex. $f_0(x) = \|x\|_1$.

$I = \operatorname{supp} x$,
I^c its complement.
x_I = restriction of x to I.
A_I = column restriction of A to I.

$\eta = A^T\lambda \in \partial f_0(x_0)$ when

(i) $\eta_I = \operatorname{sgn}(x_0)_I$
(ii) $\|\eta_{I^c}\|_\infty \le 1$.

Thm (Fuchs, 2004). x_0 is the unique minimizer of $\min \|x\|_1 : Ax = Ax_0$ when $\exists\, \eta = A^T\lambda$ such that

(i) $\eta_I = \operatorname{sgn}(x_0)_I$
(ii) $\|\eta_{I^c}\|_\infty < 1$

and furthermore A is injective on I
$(A_I v = 0 \Rightarrow v = 0)$.

Pf. Let x solve $Ax = Ax_0$.
Wish to show $\|x_0\|_1 < \|x\|_1$ unless $x = x_0$.

$$\begin{aligned} \|x_0\|_1 &= \langle \eta, x_0 \rangle && \text{because } \eta_I = \operatorname{sgn}(x_0)_I \\ &= \langle \eta, x \rangle && \text{because } \eta = A^T\lambda,\ Ax = Ax_0 \\ &= \langle \eta_I, x_I \rangle + \langle \eta_{I^c}, x_{I^c} \rangle \\ &\le \|x_I\|_1 + \|\eta_{I^c}\|_\infty \|x_{I^c}\|_1 \end{aligned}$$

if $x_{I^c} \ne 0$, then $< \|x_I\| + \|x_{I^c}\|_1 = \|x\|_1$
if $x_{I^c} = 0$, then $A_I(x_I - (x_0)_I) = 0$

$$\underset{\text{injectivity}}{\Longrightarrow} x_I = (x_0)_I \Rightarrow x = x_0 \qquad \square$$

Ideal way of building η:

$$\min_{\eta \in \operatorname{Ran} A^T} \|\eta\|_\infty \;:\; \eta_I = \operatorname{sgn}(x_0)_I$$

$\rightarrow$ no formula / no progress

Classical way of building η, although not tight:

$$\min_{\eta = A^T\lambda} \|\lambda\|_2 \;:\; \eta_I = \operatorname{sgn}(x_0)_I$$

$\Rightarrow (A^T\lambda)_I = \operatorname{sgn}(x_0)_I$
(select rows of A^T in I
$\Leftrightarrow$ select cols of A in I)

$\Rightarrow (A_I)^T\lambda = \operatorname{sgn}(x_0)_I$

$A = \boxed{\;\; | \; A_I \; | \;\;}$ underdet LS

$\Rightarrow \lambda = (A_I^T)^+ \operatorname{sgn}(x_0)_I$
(inversion OK, A_I injective)

$\Rightarrow \eta = A^T (A_I^T)^+ \operatorname{sgn}(x_0)_I$

then suffices to show $\|\eta_{I^c}\|_\infty \le 1$ or < 1

note $\eta_I = A_I^T (A_I^T)^+ \operatorname{sgn}(x_0)_I = \operatorname{sgn}(x_0)_I$.

Ex. $A_{ij} \sim N(0,1)$ iid.

means $f(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$, $\quad P(|A_{ij}| \ge u) = 2\int_u^\infty f(t)\,dt \le e^{-u^2/2}$ (\$)

Lemma 1 $X_i \sim N(0,1)$ iid

$$\mathbb{P}\left(\left|\sum c_i X_i\right| \geq u\right) \leq e^{-u^2/2\|c\|_2^2}$$

Pf. $\sum c_i X_i \sim N(0, \|c\|_2^2)$. □

Lemma 2 $A_{ij} \sim N(0,1)$ iid, $i=1,\dots,m$, $j=1,\dots,k$, $k<m$

$$\mathbb{P}\left(\sigma_{max}\left(\frac{A}{\sqrt{m}}\right) \geq 1+\sqrt{\frac{k}{m}}+t\right) \leq e^{-mt^2/2}$$

$$\mathbb{P}\left(\sigma_{min}\left(\frac{A}{\sqrt{m}}\right) \leq 1-\sqrt{\frac{k}{m}}-t\right) \leq e^{-mt^2/2}$$

Pf. Foucart, Rauhut p. 291
(1-Lipschitzness of extremal σ, conc. meas. + Gaussian width, Gordon's thm)

Now, let $A \in \mathbb{R}^{m \times n}$ # measurements = m
$A_I \in \mathbb{R}^{m \times k}$
$x_0 \in \mathbb{R}^n$ with $\|x_0\|_{\ell_0} = k$, $|I| = k$, $k<m$

$$\eta_i = \sum_{j=1}^{m} A_{ji} \Big[\underbrace{(A_I^T)^+ \operatorname{sgn}(x_0)_I}_{c}\Big]_j$$

Find conditions on m, n, k such that $|\eta_i| < 1$, $i \notin I$.

Failure probability

$$P = \mathbb{P}(\exists i \notin I : |\eta_i| \geq 1)$$

* Consider event E : $\|c\|_2 < \alpha$

then $P = \mathbb{P}(\exists i \notin I : |\eta_i| > 1 \mid E)\, \mathbb{P}(E)$

$+ \mathbb{P}(\exists i \notin I : |\eta_i| > 1 \mid E^c)\, \mathbb{P}(E^c)$.

$\leq \mathbb{P}(\exists i \notin I : |\eta_i| > 1 \mid E)$ (1)

$+ \mathbb{P}(E^c)$ (2)

- Estimate (1): fix $i \notin I$., $E = \|c\|_2 < \alpha$.

$$\mathbb{P}\left(\left|\sum_{j=1}^{m} A_{ji} c_j\right| > 1 \mid E\right) \underset{\text{lemma}}{\leq} \exp\left(-\frac{1}{2\|c\|_2^2}\right) \leq e^{-1/2\alpha^2}$$

Union bound: $n-k$ possible values of $i \notin I$,

$$\mathbb{P}\left(\exists\, i \notin I : \left|\sum_{j=1}^{m} A_{ji} c_j\right| > 1 \mid E\right) \leq (n-k)\, e^{-1/2\alpha^2} \quad (3)$$

- Estimate (2): say $\|c\|_2 \leq B$

$\|c\|_2 \geq \alpha \Rightarrow B \geq \alpha$

$\|c\|_2 \geq \alpha$ $B \geq \alpha$ $\mathbb{P}(E^c) = \mathbb{P}(\|c\|_2 \geq \alpha) \leq \mathbb{P}(B \geq \alpha)$

$$\|c\|_2 = \|(A_I^T)^{+} \operatorname{sgn}(x_0)_I\|_2 \leq (\sigma_{\min}(A_I))^{-1} \sqrt{k}$$
$$= \sigma_{\min}^{-1}(A_I/\sqrt{k})$$

$$\mathbb{P}(E^c) \leq \mathbb{P}\left(\sigma_{\min}^{-1}(A_I/\sqrt{k}) \geq \alpha\right)$$
$$= \mathbb{P}\left(\sigma_{\min}(A_I/\sqrt{m}) \leq \frac{1}{\alpha}\sqrt{\frac{k}{m}}\right)$$

lemma 2 with $1 - \sqrt{\frac{k}{m}} - t = \frac{1}{\alpha}\sqrt{\frac{k}{m}}$

$\Rightarrow t = 1 - \left(1 + \frac{1}{\alpha}\right)\sqrt{\frac{k}{m}}$

$$\leq \exp\left(-\frac{m}{2}\left(1 - \left(1+\frac{1}{\alpha}\right)\sqrt{\frac{k}{m}}\right)^2\right) \quad (4)$$

Would like $(3) + (4) \leq \underset{\text{prescribed}}{\varepsilon}$ from (3) $\leq \frac{n-1}{n}\varepsilon$

(4) $\leq \frac{\varepsilon}{n}$

$\rightarrow$ pick $e^{-1/2\alpha^2} \leq \frac{\varepsilon}{n} \Leftarrow \alpha^{-2} = 2\ln(n/\varepsilon)$

then (3) $\leq (n-k)\frac{\varepsilon}{n} \leq \frac{n-1}{n}\varepsilon$.

and (4) $\leq \frac{\varepsilon}{n}$

$\Leftrightarrow m\left(1-(1+\frac{1}{\alpha})\sqrt{\frac{k}{m}}\right)^2 \geq 2\ln(\frac{n}{\varepsilon}) = \alpha^{-2}$

$\Leftrightarrow \quad \sqrt{m} - (1+\frac{1}{\alpha})\sqrt{k} \geq \alpha^{-1}$

$\Leftrightarrow \quad \sqrt{m} \geq \alpha^{-1} + (1+\alpha^{-1})\sqrt{k}$

$\Leftarrow \quad \sqrt{m} \geq 3\alpha^{-1}\sqrt{k}$

$\Leftrightarrow \quad \boxed{m \geq 18\, k \ln(n/\varepsilon)}$

measurements $\gtrsim$ sparsity $\times \log n$

<u>Thm</u> (Donoho; "Candès-Romberg-Tao")
Let $x_0 \in \mathbb{R}^n$ be k-sparse.
$y = Ax_0$ for $A \in \mathbb{R}^{m \times n}$, $A_{ij} \sim N(0,1)$ iid
Then $\min \|x\|_1 : Ax = y$ recovers x_0
with probability $\geq 1-\varepsilon$, provided

$$m \geq C\, k \ln\left(\frac{n}{\varepsilon}\right)$$

<u>Rmk.</u> Tighter analysis (with gaussian widths) gives $m \geq 2k\ln(\frac{n}{k})$.
(Foucart, Rauhut)
The analysis presented here is tightest known for subgaussian matrices.