05/01/2018. $y = Ax_0$

$\min \|x\|_1 : Ax = y.$

<u>Thm</u>. Let $x_0 : \|x_0\|_0 = k$

$A \in \mathbb{R}^{m \times n}$, $A_{ij} \sim N(0,1)$

Recovery of x_0 w.p. $1-\varepsilon$ provided

$$m \geq Ck \ln\left(\frac{n}{\varepsilon}\right)$$

<u>Rmk</u>. Tighter analysis provides (nearly)

$$m \geq 2k \ln\left(\frac{en}{k}\right) + \ln \varepsilon^{-1} \simeq 2k \ln\left(\frac{en}{k}\right)$$

Proof technique: dual certification.

<u>Thm</u> Recovery of x_0 when

$\exists\ \eta = A^T\lambda$:

(i) $\eta_I = \mathrm{sgn}(x_0)_I$

(ii) $\|\eta_{I^c}\|_\infty < 1$

and A_I injective $\quad (A_I v = 0 \Rightarrow v = 0)$

Inexact + robust version (to noise & nonstrict sparsity)

<u>Def</u>. (ℓ_p error of best k-term approximation)

for $p > 0$, $x \in \mathbb{R}^n$,

$\sigma_k(x)_p := \inf \|x - y\|_p$, y is k-sparse}

(inf realized by picking k |largest| components of x)

Setting : $y = Ax_0 + e$ with

- $\|e\|_2 \le \varepsilon$
- injectivity of A only controlled on I with $|I| = k$

then, at best, any recovered x obeys
$$\|x - x_0\| \lesssim \varepsilon + \sigma_S(x_0)_2$$

Not. $A = [a_1 | \cdots | a_n]$, $x = \operatorname{argmin} \|x\|_1 : \|Ax - y\| \le \varepsilon$

Thm. Let $x_0 \in \mathbb{R}^n$ with k largest components on I. Let $y = Ax_0 + e$ with $\|e\|_2 \le \varepsilon$ Assume $\exists\ \delta, \beta, \gamma, \theta, \tau$ ($\delta < 1$) such that

- $\|A_I^* A_I - I\|_{2\to 2} \le \delta$ (need $\delta < 1$)
- $\max_{i \notin I} \|A_I^* a_i\|_2 \le \beta$ (think $\beta = 1$)

and $\exists\ \eta = A^T \lambda$ such that

- $\|\eta_I - \operatorname{sgn}(x_0)_I\|_2 \le \gamma$
- $\|\eta_{I^c}\|_\infty \le \theta$ (need $\theta < 1$)
- $\|\lambda\|_2 \le \tau \sqrt{k}$

If $p = \theta + \frac{\beta\gamma}{1-\delta}$, then

$$\|x - x_0\|_2 \le C_1 \sigma_k(x)_1 + C_2 \sqrt{k}\, \varepsilon,$$

(where C_1, C_2 depend on $\delta, \beta, \gamma, \theta, \tau$)

Rmk: η is still "about x_0", even though the minimizer is no longer x_0
$\rightarrow$ not exactly same interpretation as before

Dual certification for other problems.

Low-rank matrix recovery :

$$\min \underbrace{\|X\|_*}_{\sum \sigma_i(X)} : \underbrace{A}_{\text{linear}}(X) = y \quad , \quad X \in \mathbb{R}^{m \times n}, \; y \in \mathbb{R}^p$$

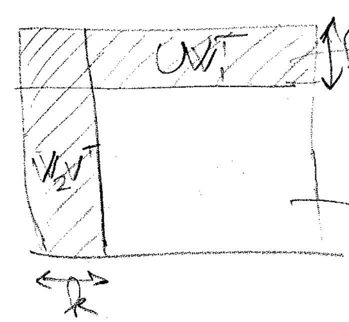
Notion to replace I (support) :
for $X = U\Sigma V^T \in \mathbb{R}^{m\times n}$, let

$$T = \{ M \in \mathbb{R}^{m\times n} : M = U W_1^T + W_2 V^T \}$$

$$T^\perp = \{ N \in \mathbb{R}^{m \times n} : \langle M, N \rangle = 0 \;\; \forall M \in T \}$$

ex. $U = \left[e_1 \middle| \cdots \middle| e_k \right]$ $\quad V^T = \begin{bmatrix} e_1^T \\ \vdots \\ e_k^T \end{bmatrix}$

unit canonical in $\mathbb{R}^m$

$M =$ 

$\to$ matrices supported in the shaded region are in T

$\to T^\perp$ ("weak orthogonality")

$$f(X) = \|X\|_1 : \qquad (X = U \Sigma V^T)$$

$$\partial f(X) = \{ UV^T + W : \|W\| \le 1, \; U^T W = 0, \; W V = 0 \}$$

$$Y = \begin{bmatrix} UV^T & 0 \\ 0 & W \end{bmatrix}$$

because $U^T W = 0$

$\|W\| \le 1$

("strong orthogonality")

because $WV = 0$

so $Y \in \partial f(X)$ means
- $\underbrace{Y_T}_{\text{proj. onto } T} = UV^T$ (1)
- $\|Y_{T^\perp}\| \le 1$. (2)

<u>Def</u>. $A^* : \mathbb{R}^p \to \mathbb{R}^{m \times n}$ defined from

$$\langle A(X), y \rangle_{\mathbb{R}^p} = \langle X, A^*(y) \rangle_{\mathbb{R}^{m \times n}} \quad \forall X, y$$

<u>Thm</u>. If $\exists\, Y \in \operatorname{Ran} A^*$: (1), (2) hold, then X_0 is a minimizer of $\min \|X\|_* : A(X) = y = A(X_0)$

<u>Thm</u>. (Candès, Recht, 2008). If in addition, $\|Y_{T^\perp}\| < 1$, and A is injective on T ($\forall M \in T$, $A(M) = 0 \Rightarrow M = 0$), then X_0 is the unique minimizer of $\min \|X\|_* : A(X) = y = A(X_0)$

<u>Pf</u>. Wish to show $\|X\|_* > \|X_0\|_*$ unless $X = X_0$. Let $Z \in \partial f(X_0)$, unspecified (not Y)

$$\|X\|_* \ge \|X_0\|_* + \langle Z, X - X_0 \rangle$$
$$\ge \|X_0\|_* + \langle Z - Y, X - X_0 \rangle$$

because $A(X - X_0) = 0$

Pick $Z_T = Y_T$,
$Z_{T^\perp}$ such that $\langle Z_{T^\perp}, X \rangle = \|X_{T^\perp}\|_*$
and $\|Z_{T^\perp}\| \le 1$.

(do this by $X_{T^\perp} = U'\Sigma' V'^T$
then $Z_{T^\perp} = U'V'^T$, so

$$\langle Z_{T^\perp}, X\rangle = \langle Z_{T^\perp}, X_{T^\perp}\rangle$$
$$= \langle U'V'^T, U'\Sigma' V'^T\rangle$$
$$= \operatorname{tr}\langle V'U'^T U'\Sigma' V'^T\rangle$$
$$= \operatorname{tr}\Sigma' = \|X_{T^\perp}\|_*$$

$\|Z_{T^\perp}\| = 1$)

then

$$\|X\|_* \geq \|X_0\|_* + \langle Z_{T^\perp} - Y_{T^\perp}, X_{T^\perp}\rangle$$

$$\geq \|X_0\|_* + (\underbrace{1 - \|Y_{T^\perp}\|}_{>0})\|X_{T^\perp}\|_*$$

$$> \|X_0\|_* \quad \text{when } X_{T^\perp} \neq 0$$

When $X = X_T$, then $A(X_T - X_0) = 0$
$\Rightarrow X_T = X_0$
$\Rightarrow X = X_0$. $\square$

(l_1 argument would not carry over because $\|X_T\|_* + \|X_{T^\perp}\|_* \not\leq \|X\|_*$).

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

$\sqrt{5}$ $\quad$ 1 $\quad$ 2