E9 203 Compressed Sensing & Sparse SP

19 Apr. 2021

Last time

- Recovery of individual sparse vecs. via tangent cones to ℓ_1 balls

Real-valued case:

$x \in \mathbb{R}^N$, convex cone:

$T(x) = \operatorname{cone}\left(\{z - x : z \in \mathbb{R}^N, \|z\|_1 \le \|x\|_1\}\right)$

Cone (T): smallest convex cone containing T

$\operatorname{cone}(T) = \{\sum_{j=1}^{n} t_j x_j : n \ge 1,\ t_1 \dots t_n \ge 0,\ x_1 \dots x_n \in T\}$

Thm. 4.35 $A \in \mathbb{R}^{m\times N}$, $x \in \mathbb{R}^N$ is the unique min. of $\|z\|_1$ s.t. $Az = Ax$ iff $\mathcal{N}(A) \cap T(x) = \{0\}$.

Thm. 4.36 $A \in \mathbb{R}^{m\times N}$, $x \in \mathbb{R}^N$, $y = Ax + e \in \mathbb{R}^m$ with $\|e\|_2 \le \eta$. If

$$\inf_{\substack{v \in T(x)\\ \|v\|_2 = 1}} \|Av\|_2 \ge \tau$$

for some $\tau > 0$, then a minimizer $x^\#$ of $\|z\|_1$ s.t. $\|Az - y\|_2 \le \eta$ satisfies $\|x - x^\#\|_2 \le \frac{2\eta}{\tau}$.

Today:

- Low rank matrix recovery
- Coherence.

First, one clarification. We saw:

Thm. 4.26: Given $A \in \mathbb{C}^{m\times N}$, the vec. $x \in \mathbb{C}^N$ with $\operatorname{supp}(x) = S$ is the unique minimizer of $\|z\|_1$ s.t. $Az = Ax$ if one of the foll. equivalent conditions hold:

(a) $\left|\sum_{j\in S} \overline{\operatorname{sgn}(x_j)}\, v_j\right| < \|v_{\bar{S}}\|_1 \quad \forall v \in \mathcal{N}(A)\setminus\{0\}$.

(b) A_S is injective & $\exists\ h \in \mathbb{C}^m$ s.t.

$$(A^H h)_j = \operatorname{sgn}(x_j), \quad j \in S$$

$$|(A^H h)_\ell| < 1, \quad \ell \in \bar{S}.$$

Proof: Step 1: (a) $\Rightarrow$ unique recovery of x

Step 2: (b) $\Rightarrow$ (a).

Step 3: (a) $\Rightarrow$ (b).

For step 3, we defined

$\mathcal{D} \triangleq \{z \in \mathbb{C}^N : Az = Ax\}$

Using the result on separating hyperplanes for convex sets, $\exists\ w \in \mathbb{C}^N$ s.t.

$$\mathcal{D} \subset \{z \in \mathbb{C}^N : \operatorname{Re}\langle z, w\rangle = \|x\|_1\} \quad \text{—} (*)$$

Then we showed that $\|w_{\bar{S}}\|_\infty \le \nu < 1$

$\operatorname{Re}\langle x, w\rangle = \|x\|_1 \Rightarrow w_j = \operatorname{sgn}(x_j)\ \forall j \in S$

and we claimed $\operatorname{Re}\langle v, w\rangle = 0\ \forall v \in \mathcal{N}(A)$. Why?

$v \in \mathcal{N}(A) \Rightarrow x - v \in \mathcal{D}$.

$\Rightarrow$ by $(*)$, $\operatorname{Re}\langle x - v, w\rangle = \|x\|_1$

$\Rightarrow \underbrace{\operatorname{Re}\langle x, w\rangle}_{= \|x\|_1} - \operatorname{Re}\langle v, w\rangle = \|x\|_1$

$\Rightarrow \operatorname{Re}\langle v, w\rangle = 0 \quad \forall v \in \mathcal{N}(A)$.

Then we concluded the proof by saying $w \in \mathcal{N}(A)^\perp = \mathcal{R}(A^H)$

$\Rightarrow w = A^H h$ for some $h \in \mathbb{C}^m$. ☒

Low rank matrix recovery:

$X \in \mathbb{C}^{n_1\times n_2}$, $\operatorname{rank}(X) \le r$

$y = \mathcal{A}(X) \in \mathbb{C}^m$

$\mathcal{A}$ is a linear map $\mathbb{C}^{n_1\times n_2} \longrightarrow \mathbb{C}^m$.

Equivalent of the (P_0) problem:

$$\min_{Z \in \mathbb{C}^{n_1\times n_2}} \operatorname{rank}(Z) \text{ s.t. } \mathcal{A}(Z) = y.$$

NP hard (see Ex. 2.11)

$\operatorname{rank}(Z) = \|\underline{\sigma}(Z)\|_0$, where $\underline{\sigma}(Z) = \begin{bmatrix}\sigma_1(Z)\\ \vdots \\ \sigma_n(Z)\end{bmatrix}$,

$n = \min(n_1, n_2)$, is the vec. containing the singular values of Z.

Relax (P_0) as follows:

$$(P_*) \quad \min_{Z \in \mathbb{C}^{n_1\times n_2}} \|Z\|_* \text{ s.t. } \mathcal{A}(Z) = y.$$

$\|\cdot\|_*$ is the nuclear norm:

$$\|Z\|_* = \sum_{j=1}^{n} \sigma_j(Z), \quad n = \min\{n_1, n_2\}.$$

Properties of $\|\cdot\|_*$:

1. $\|\cdot\|_*$ is indeed a norm (Appendix (A.25))
2. (P_*) is a convex opt. problem
3. Equivalent to an SDP.

Thm. 4.40

Given a linear map $\mathcal{A}: \mathbb{C}^{n_1\times n_2} \longrightarrow \mathbb{C}^m$, every matrix $X \in \mathbb{C}^{n_1\times n_2}$ with rank at most r is the unique soln of (P_*) with $y = \mathcal{A}(X)$ iff $\forall\ M \in \mathcal{N}(\mathcal{A})\setminus\{\underline{0}\}$ with singular values $\sigma_1(M) \ge \sigma_2(M) \ge \dots \ge \sigma_n(M) \ge 0$, $n \triangleq \min\{n_1, n_2\}$,

$$\sum_{j=1}^{r} \sigma_j(M) < \sum_{j=r+1}^{n} \sigma_j(M) \quad [\text{Rank NSP}].$$

Proof: Suppose every $X \in \mathbb{C}^{n_1\times n_2}$ of rank $\le r$ is the unique soln. to (P_*) with $y = \mathcal{A}(X)$.

Consider the SVD of $M \in \mathcal{N}(\mathcal{A})\setminus\{0\}$

$$M = \underbrace{U}_{\mathbb{C}^{n_1\times n_1}} \begin{pmatrix}\sigma_1 & & 0\\ & \ddots & \\ 0 & & \sigma_n\end{pmatrix} \underbrace{V^H}_{\mathbb{C}^{n_2\times n_2}},$$

$\sigma_1 \ge \sigma_2 \ge \dots \ge \sigma_n \ge 0$.

Let $M_1 = U \begin{pmatrix}\sigma_1 & & & & 0\\ & \ddots & & & \\ & & \sigma_r & & \\ & & & 0 & \\ 0 & & & & \ddots\ 0\end{pmatrix} V^H$

$M_2 = U \begin{pmatrix}0 & & & & 0\\ & \ddots & & & \\ & & 0 & & \\ & & & -\sigma_{r+1} & \\ 0 & & & & \ddots\ -\sigma_n\end{pmatrix} V^H$

$\Rightarrow M = M_1 - M_2$, $\mathcal{A}(M) = 0 \Rightarrow \mathcal{A}(M_1) = \mathcal{A}(M_2)$

Since $\operatorname{rank}(M_1) \le r$, its nuclear norm must smaller than that of M_2, i.e.,

$$\sigma_1 + \dots + \sigma_r < \sigma_{r+1} + \dots + \sigma_n, \text{ as desired.}$$

Conversely, suppose

$$\sum_{j=1}^{r} \sigma_j(M) < \sum_{j=r+1}^{n} \sigma_j(M) \quad \forall M \in \mathcal{N}(\mathcal{A})\setminus\{0\}$$

with $\sigma_1(M) \ge \dots \ge \sigma_n(M) \ge 0$ $\quad (n = \min\{n_1, n_2\})$.

Consider $X \in \mathbb{C}^{n_1\times n_2}$, $\operatorname{rank}(X) \le r$

$Z \in \mathbb{C}^{n_1\times n_2}$, $Z \ne X$, $\mathcal{A}(Z) = \mathcal{A}(X)$

Want to show that $\|Z\|_* > \|X\|_*$.

Set $M = X - Z \in \mathcal{N}(\mathcal{A})\setminus\{0\}$.

Result: Lemma A.18:

$X \in \mathbb{C}^{n_1\times n_2}$, $Z \in \mathbb{C}^{n_1\times n_2}$, $n = \min\{n_1, n_2\}$

for any $k \in [n]$,

$$\sum_{j=1}^{k} |\sigma_j(X) - \sigma_j(Z)| \leq \sum_{j=1}^{k} \sigma_j(X-Z) .$$

$\sigma_j(M)$, $\sigma_j(Z)$, $\sigma_j(X)$ satisfy

$$\|Z\|_* = \sum_{j=1}^{n} \sigma_j(X-M) \geqslant \sum_{j=1}^{n} |\sigma_j(X) - \sigma_j(M)|$$

Now, for $j \in [r]$,

$$|\sigma_j(X) - \sigma_j(M)| \geqslant \sigma_j(X) - \sigma_j(M)$$

and for $r+1 \leq j \leq n$

$$|\sigma_j(X) - \sigma_j(M)| = \sigma_j(M)$$

$$\|Z\|_* \geqslant \sum_{j=1}^{r} \sigma_j(X) - \sigma_j(M) + \sum_{j=r+1}^{n} \sigma_j(M)$$

$$= \sum_{j=1}^{r} \sigma_j(X) + \underbrace{\left[\sum_{j=r+1}^{n} \sigma_j(M) - \sum_{j=1}^{r} \sigma_j(M) \right]}_{+\text{ve}, > 0}$$

$$> \sum_{j=1}^{r} \sigma_j(X) = \|X\|_* . \qquad \boxtimes$$

See Ex. 4.19 & 4.20 for stable and robust versions of the rank NSP.

Chapter 5 : Coherence

- Simple measure of the "quality" of a matrix
- Smaller the better.

In this chapter, assume that the cols of $A \in \mathbb{C}^{m \times N}$ have unit ℓ_2 norm ($\|a_i\|_2 = 1$, $i \in [N]$)

Defn. $A \in \mathbb{C}^{m \times N}$ with ℓ_2 normalized cols $a_1 \dots a_N$ s.t. $\|a_i\|_2 = 1$, $i \in [N]$.
The coherence $\mu(A)$ is defined as

$$\mu \triangleq \mu(A) \triangleq \max_{1 \leq i \neq j \leq N} |\langle a_i, a_j \rangle| .$$