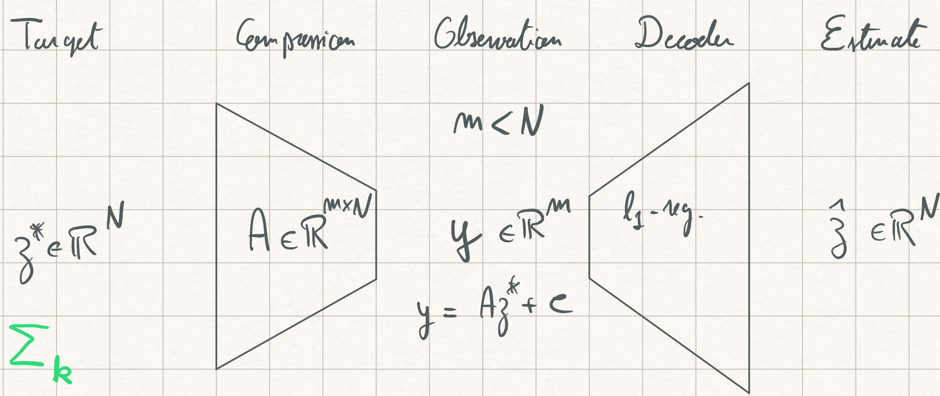
Lesson 4

The big picture:



⊗ Conditions on A

- ERC_k: $\forall S \subset [N]$, $\#S = k$, $\quad \| (A_S^+) A_{S^c} \|_{1 \to 1} < 1$

  where $A_S^+ = (A_S^T A_S)^{-1} A_S^T$ and A normalized ($\|A_i\|_2^2 = 1$)

$\iff$ $\| A_S^T A_S u \|_\infty > \| A_{S^c}^T A_S u \|_\infty \quad \forall u \neq 0.$

A_S injective.

- NSP_1: relative to a set $S \subseteq [N]$: $\forall v \in \mathbb{R}^N$ s.t $v \in \ker A \setminus \{0\}$

$$\|v_S\|_1 < \|v_{S^c}\|_1$$

- RIP_k: $\exists\, \delta_k > 0$ s.t $\forall x \in \Sigma_k$,

$$(1 - \delta_k)\, \|x\|_2^2 \leq \|Ax\|_2^2 \leq (1 - \delta_k)\, \|x\|_2^2$$

$\to$ A normalized: $\|A_i\|_2^2 = 1$

- Coherence Property

Def: A with normalized columns

$$\mu = \mu(A) = \max_{1 \leq i \neq j \leq N} |\langle A_i, A_j \rangle|$$

Rk: A normalized

$$A^T A - Id_N = \left(\delta_{i \neq j} \langle A_i, A_j \rangle \right)_{i,j}$$

Def: ℓ_1-coherence function

$$\mu_1(k) = \max_{i \in [N]} \max \left\{ \sum_{j \in S} |\langle A_i, A_j \rangle| , \; S \subset [N], \; \# S = k, \; i \notin S \right\}$$

$k \in [N-1]$

Note that $\forall\, k \in [N-1]$,

$$\boxed{\mu \leq \mu_1(k) \leq k\mu}$$

Theorem (Coherence - RIP)

- A normalized
- $\forall\, x \in \Sigma_k$,

  $2k$

$$\boxed{\begin{aligned}(1-\mu_1(k-1))\,\|x\|_2^2 \leq \|Ax\|_2^2 &\leq \cdots \\ \cdots\ (1+\mu_1(k-1))\,\|x\|_2^2 &\end{aligned}}$$

- If $\mu_1(k-1) < 1$ Then RIP_k holds

  $2k-1$ $\quad$ $2k$

__Proof:__ • Let $S \subset [N]$ be s.t $\# S = k$

• Let $x \in \Sigma_k$ be s.t $\operatorname{Supp} x = S$

$$\|Ax\|_2^2 = x_S^T A_S^T A_S x_S$$

(indeed $Ax = A_S x_S$)

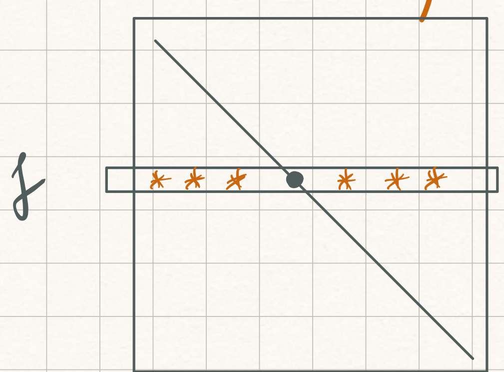
$$\lambda_{max} = \max_{\substack{x \in \mathbb{R}^N \\ \operatorname{Supp} x \subset S \\ \|x\|_2 = 1}} \{ \langle A_S^T A_S x_S , x_S \rangle \}$$

$$\lambda_{min} = \min_{\substack{x \in \mathbb{R}^N \\ \operatorname{Supp} x \subset S \\ \|x\|_2 = 1}} \{ \langle A_S^T A_S x_S , x_S \rangle \}$$

__Note that:__

• $(A_S^T A_S)_{ii} = 1$

$$\cdot \; r_j = \sum_{l \in S,\, l \neq j} \underbrace{\left| (A_S^T A_S)_{j,l} \right|}_{*}$$



$$= \; A_S^T A_S$$

$$r_j = \sum_{l \in S,\, l \neq j} \left| \langle A_l, A_j \rangle \right| \leq \mu_1 (k-1)$$

$$\forall j \in S$$

<u>By Gershgorin's disk theorem</u>

$$|\text{eigenvalue}| \in [1 - \mu_1(k-1),\; 1 + \mu_1(k-1)]$$

eigenvalues $\subset [1-\mu_1(k-1), 1+\mu_1(k-1)]$

Theorem A is normalized

If $\underline{\mu_1(k) + \mu_1(k-1) < 1}$ Coherence Condition CC_k

then ERC_k holds

Rk: $\cdot\ \mu_1(2k-1) \leq \mu_1(k) + \mu_1(k-1)$

$CC_k \Rightarrow \mu_1(2k-1) < 1 \Rightarrow RIP_{2k}$

$\cdot\ CC_k \Rightarrow ERC_k$

Proof: $CC_k \Rightarrow ERC_k$

We need to prove that:

$\cdot\ A_S$ injective $\Big)$ implied by $\mu_1(2k-1) < 1 \Rightarrow RIP_{2k}$

- $\|A_S^T r\|_\infty > \|A_{S^c}^T r\|_\infty$

where $r = A_S z$ with $z \neq 0$

- Let $z \neq 0$, set $r = A_S z$, and chose:

$$\ell \in S \quad \text{s.t} \quad |z_\ell| = \|z\|_\infty$$

- Note that for $j \in S^c$

$$\begin{aligned}
|\langle A_j, r\rangle| &= \Big|\sum_{i \in S} z_i \langle A_i, A_j\rangle\Big| \\
&\leq \sum_{i \in S} |z_i|\, |\langle A_i, A_j\rangle| \\
&\leq |z_\ell| \quad \mu_1(k) \\
&\lneqq
\end{aligned}$$

- And for $i \in S$,

$$|\langle A_\ell, r\rangle| = \Big|\sum_{j \in S} z_j \langle A_\ell, A_j\rangle\Big|$$

$$\geq |z_\ell|\,|\langle A_\ell, A_\ell\rangle| - \sum_{\substack{j\neq \ell\\ j\in S}} |z_j|\,|\langle A_j, A_\ell\rangle|$$

$$\geq |z_\ell| - |z_\ell|\,\mu_1(k-1)$$

$$= |z_\ell|\,(\underbrace{1-\mu_1(k-1)}_{>\mu_1(k)})$$

$$\|A_S^T r\|_\infty \geq |\langle A_\ell, r\rangle| \geq |z_\ell|\,(1-\mu_1(k-1))$$

$$> |z_\ell|\,\mu_1(k)$$

$$\geq \|A_{S^c}^T r\|_\infty \qquad \square$$

$$CC_k \Rightarrow RIP_k$$

$$CC_k \Rightarrow ERC_k$$

II) Stability and robustness

<u>Stability</u> : z^* is not sparse

<u>Robustness</u> : e is not zero

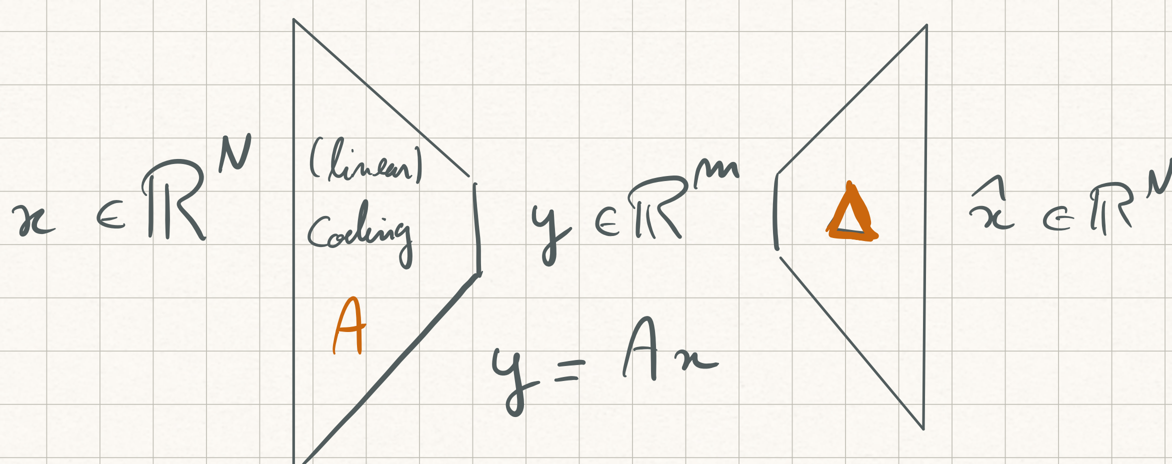
Instance Optimality

$$\boxed{\sigma_k(x)_p := \inf \left\{ \|x - z\|_p : z \in \Sigma_k \right\}}$$

- $k \in [N]$ sparsity
- $p \in [1, \infty]$ ℓ_p-norm

<u>Rk</u>: $x \in \Sigma_k$, $\sigma_k(x)_p = 0$

Coding - Decoding Scheme



$\Delta : \mathbb{R}^m \longrightarrow \mathbb{R}^N$ <u>decoder</u>

<u>Def</u> l_p- instance optimality <u>of order k</u>

We say that Δ is

l_p- instance optimal of order k

<u>iff</u> $\exists C > 0, \forall x \in \mathbb{R}^N,$

$$\| x - \underbrace{\Delta(Ax)} \|_p \leq C \, \sigma_k(x)_p$$

<u>Theorem</u> Let $A \in \mathbb{R}^{m \times N}$ be given.

<u>If</u> $\exists \Delta$ s.t (A, Δ) is l_1- instance optimal of order k with constant $C > 0$

<u>then</u>

$\forall v \in \ker A,$

$$\|v\|_1 \leq C \, \sigma_{2k}(v) \qquad (**)$$

($C = 2$ is the NSP_1)

Conversely if $(**)$ holds

then $\exists \Delta$ s.t (A, Δ) is l_1-instance optimal of order k and constant $2C$.

Proof: Let $v \in \ker A$
Let S be an index of the k largest entries of v

Inst. Opt $\Rightarrow$ $-v_S = \Delta(\underbrace{A(-v_S)}_{A v_{S^c}}) = \Delta(A v_{S^c})$

$$\|v\|_1 = \|v_S + v_{S^c}\|_1 = \|v_{S^c} - \Delta(A v_{S^c})\|_1$$

$$\leq C \underbrace{\sigma_k(v_{S^c})_1}_{\text{sum on the } N-2k \text{ least abs. val coeff}} = C\,\sigma_{2k}(v)_1$$

Conversely, Assume (**) condition.

Define Δ as:

$$\boxed{\underbrace{\Delta(y)}_{z^*} = \arg\min_{\substack{z: \\ Az = y}} \left\{ \sigma_k(z)_1 \right\}}$$

$y = Ax$
x is feasible

Let $x \in \mathbb{R}^N$, (**) with $v = x - \Delta(Ax) \in \ker A$

$$\left(Av = Ax - \underbrace{A(\underbrace{\Delta(Ax)}_{z^*})}_{Ax} = 0 \right)$$

$$\|x - \Delta(Ax)\|_2 \leq C\,\sigma_{2k}(x - \Delta(Ax))_1$$

$$\leq C \left[\sigma_k(x)_1 + \underbrace{\sigma_k(\Delta(Ax))_1}_{\leq \sigma_k(x)_1 \quad (x \text{ feasible})} \right]$$

$$\leq 2C\, \sigma_k(x)_1 \qquad \square$$

<u>Theorem</u> <u>If</u> (A, Δ) are ℓ_1-instance optimal of order k and constant $C > 0$

<u>then</u>

$$\boxed{m \geq c\, k \ln(eN/k)}$$

where c depends only on C.

<u>Def</u> <u>Stable NSP</u> with constant

$$0 < \rho < 1$$

iff $\forall v \in \ker A, \ \|v_S\|_1 \leq \rho \|v_{\bar{S}}\|_1$
$\forall S \subset [N]$ s.t. $\# S = k$,

Theorem **Assume** A satisfies **Stable NSP** of order k and constant $0<\rho<1$

Then $\forall x \in \mathbb{R}^N$, any solution $\hat{x}$ of **Basis Pursuit**:

$$\hat{x} \in \arg\min_{Az = Ax} \|z\|_1$$

is such that

$$\boxed{\|x-\hat{x}\|_1 \leq \frac{2(1+\rho)}{1-\rho} \sigma_k(x)_1}$$