Inverse problems and sparsity

Today : - Inverse problems

- sparsity

- Identifiability

- Algorithms

- Recovery guarantees

- Compressed Sensing

1 Inverse problems

- $x \in \mathbb{R}^N$ unknown

  $m < N$

- Sparse model :

  - simplest model :

    x is s-sparse if

    $\|x\|_0 := \#\{ i \in [N] : x_i \neq 0 \} \leq s$

- sparsity in a basis

$x = B z$ with z is sparse

and B $N \times N$ matrix

- sparsity in a dictionary

$x = D z$ with z is sparse

and D $N \times K$ matrix

$K \geq N$

example : $D = [I_N \ \ F_N]$

- Inverse Problems $y = A x$ with A $m \times N$ matrix

$$= \left| \begin{array}{ll} A\,x & x \text{ sparse} \\ A\,B\,z & z \text{ sparse} \\ A\,D\,z & z \text{ sparse} \end{array} \right.$$

For ease of notation : $A \leftarrow A, AB$ or AD

$z^* \leftarrow x, z$ or z

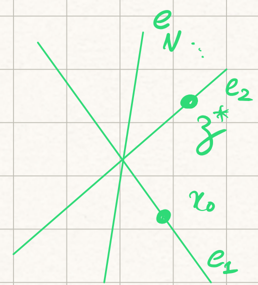
$$\boxed{y = A z^*}$$

$m \times N$ matrix A

N vector z

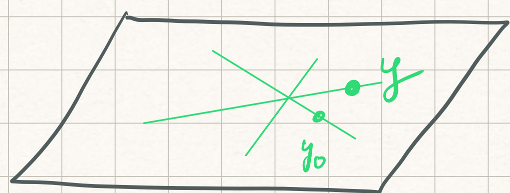
Geometric insight $\quad y = Az^* \quad z^*$ 1-sparse $(\|z\|_0 = 1)$

$\mathbb{R}^N$



feasible set
= finite union of low-dimensional subspaces
$(\cup_\circ S)$

$\mathbb{R}^m$



<u>Def</u>: $z^* \in \Sigma_k$ is <u>identifiable</u> from y s.t $y = Az^*$

iif $\forall z \in \mathbb{R}^N$,

$$\begin{pmatrix} y = Az^* \\ \text{and } z \in \Sigma_k \end{pmatrix} \Rightarrow (z = z^*)$$

where $\Sigma_k = \{ z \in \mathbb{R}^N : \|z\|_0 \leq k \}$

<u>Question</u>: Given A and k, when can we be sure that every k-sparse vector is identifiable?

Answer: This property is equivalent to

$$\boxed{\ker A \cap \Sigma_{2k} = \{0\}}$$

proof: exercice.

② Study of the l_0 - program:

$$(P_0) \qquad \min_z \|z\|_0 \quad \text{s.t} \quad y = A z^*$$

Prop: Given A and k we have (i) $\Leftrightarrow$ (ii) with

i) $\forall z^* \in \Sigma_k$, z^* unique solution to (P_0) with observation $y = Az^*$

ii) $\ker A \cap \Sigma_{2k} = \{0\}$

proof: exercice

Remark: (P_0) is NP-hard. (admitted)

Noisy setting: $y = Az^* + e$ with $e \in \mathbb{R}^m$
a additive error term
(one may assume that $\|e\|_2 \leq \varepsilon$)

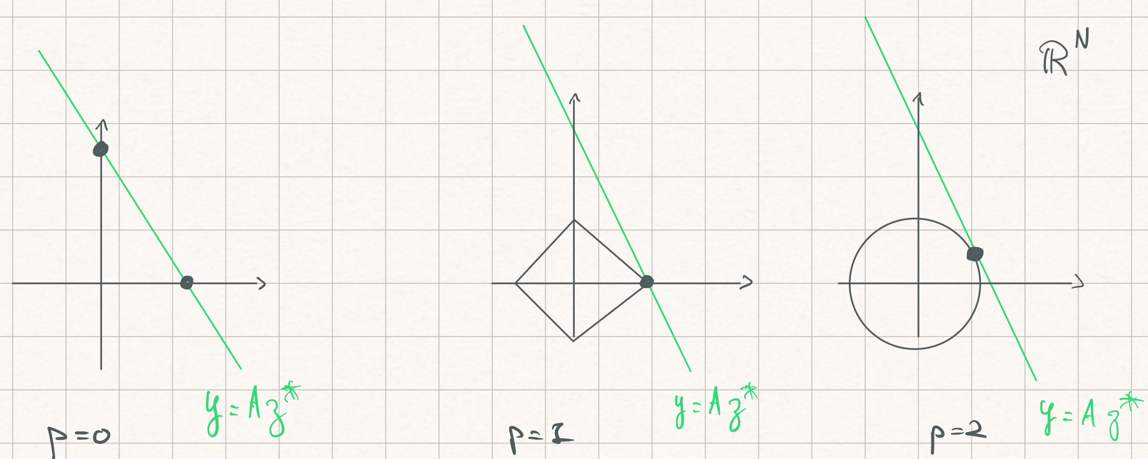
Noisy versions of (P_0):

$$\begin{cases} \min \|z\|_0 \quad \text{s.t} \quad \|y - Az\|_2 \leq \varepsilon \\ \min \|y - Az\|_2 \quad \text{s.t.} \quad \|z\|_0 \leq k \\ \min \frac{1}{2}\|y - Az\|_2^2 + \lambda \|z\|_0 \end{cases}$$

Remark: They are NP-hard (admitted)

③ Convex relaxation of the l_0 penalty.

On a display



(P_1) $\min\limits_{z} \|z\|_1$ s.t. $y = Az^*$

Theorem There exists a m-sparse solution to (P_1).

proof: $\mathcal{E} = A B_1$ symmetric convex polytope

$$B_1 = \{ z \in \mathbb{R}^N : \|z\|_1 \leq 1 \}$$

with faces of dimension $0, 1, \dots, N-1$ (affine dimension)

$$\mathcal{F}_k = \bigcup_{\substack{S \subseteq [N] \\ |S| = k+1}} \bigcup_{\varepsilon \in \{-1,+1\}^{k+1}} \left\{ z \in \mathbb{R}^N : \sum_{i=1}^{k+1} \varepsilon_i \alpha_i e_{s_i}, \ \sum \alpha_i = 1, \ \alpha_i \geq 0, \ S = \{s_1, \dots, s_k\} \right\}$$

$$|\mathcal{F}_k| = \binom{N}{k+1} 2^{k+1}$$

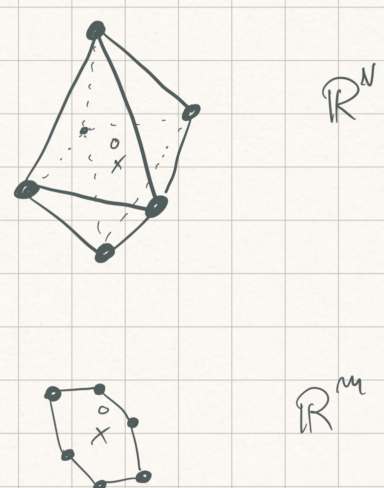
We deduce that $\mathcal{E}$ has faces of dimension $0, 1, \dots, m-1$

$$\mathcal{G}_k \subseteq A \mathcal{F}_k$$

Note that

$$y \in \|\hat{z}\|_1 A B_1$$

where $\hat{z}$ any solution to (P_1)



In particular, y belongs to a face of $\|\hat{z}\|_1 \mathcal{E}$.
By the theorem of Caratheodory y can be expressed as an affine combination of at most m points of $\mathcal{G}_0$ (extrem points of $\|\hat{z}\|_1 \mathcal{E}$) $\square$

<u>Remark</u>: We will prove that z^* is the unique solution to (P_1) under the "Null-Space Property".

④ <u>Algorithms</u> • $BP \hookrightarrow LP$.

•
$$\min_{z \in \mathbb{R}^N} F(z) + R(z) \qquad (1)$$

with F convex differentiable and R convex
$\left(F(z) = \frac{1}{2}\|y - Az\|_2^2\right)$ $\quad$ $(R(z) = \lambda\|z\|_1)$

then $0 \in \nabla F(\hat{z}) + \partial R(\hat{z})$

$$0 \in \gamma \nabla F(\hat{z}) + \partial(\gamma R)(\hat{z})$$

$$0 \in \hat{z} - (\hat{z} - \gamma \nabla F(\hat{z})) + \partial(\gamma R)(\hat{z})$$

$$\hat{z} \in \underbrace{\operatorname*{argmin}_{z} \left\{ \frac{1}{2}\|z - (\hat{z} - \gamma\nabla F(\hat{z}))\|_2^2 + \gamma R(z) \right\}}_{=: \, \mathrm{prox}_{\gamma R}(\hat{z} - \gamma \nabla F(\hat{z}))}$$

$$\Leftrightarrow \quad \boxed{\hat{z} \in \operatorname{prox}_{\gamma R}\left(\hat{z} - \gamma \nabla F(\hat{z})\right)} \quad \text{(fixed point)}$$

$$\operatorname{prox}_{\gamma R}(x) = \arg\min_u \left\{ \frac{1}{2} \| x - u \|_2^2 + \gamma R(u) \right\}$$

$$= \left\{ u \ : 0 \in u - x + \partial(\gamma R)(u) \right\}$$

<u>exercice</u> Prove that

$$\left(\operatorname{prox}_{\gamma \|\cdot\|_1}(x)\right)_i = \left(S_\gamma(x)\right)_i := \begin{cases} x_i - \gamma & \text{if } x_i \geq \gamma \\ 0 & \text{if } x_i \in]-\gamma, \gamma[\\ x_i + \gamma & \text{if } x_i \leq -\gamma \end{cases}$$

"Soft thresholding"

and deduce that any solution to

$$(P_1) \qquad \min \left\{ \frac{1}{2} \| y - Az \|_2^2 + \lambda \| z \|_1 \right\}$$

satisfies $\hat{z} = S_{\gamma\lambda}\left(\hat{z} - \gamma A^T(A\hat{z} - y)\right)$

- Dual program of (P_1) :

$(\mathcal{D}_1)$ $\min \|y - t\|_2^2$ s.t. $\|A^T t\|_\infty \leq \lambda$

indeed $\mathcal{L}(z, y'; t)$

$$= \frac{1}{2}\|y - y'\|_2^2 + \lambda\|z\|_1 + \langle Az - y', t\rangle$$

$$\sup_t \mathcal{L} = \frac{1}{2}\|y - y'\|_2^2 + \lambda\|z\|_1 + \chi_{y' = Az}$$

$$\inf_{z, y'} \mathcal{L} = \underbrace{\frac{1}{2}\|y - y'\|_2^2 - \langle y', t\rangle}_{\nabla = 0 \iff y' - y - t = 0} + \underbrace{\lambda\|z\|_1 + \langle z, A^T t\rangle}_{\chi_{\|A^T t\|_\infty \leq \lambda}}$$

$$= -\frac{1}{2}\|y-t\|_2^2 + \frac{1}{2}\|y\|_2^2$$

Primal dual relation: $\bullet\ t = Az - y$

$\bullet\ \lambda \|z\|_1 = \langle z, A^T t\rangle$

$$\rightarrow \boxed{\frac{A^T(Az-y)}{\lambda} = \operatorname{sign}(z)}$$

Greedy methods: OMP

$$S^{n+1} = S^n \cup \{j_{n+1}\}$$

$$j_{n+1} = \arg\max \{ |A^T(y - Ax^n)|_j \}$$