Non-convex inverse problems

March 16, 2026, 2 hours

You can use any written or printed material.
Note that the subject is long, but the total number of points (not including bonus questions) is above 20.
The bonus questions are either not directly related to inverse problems or difficult. They will not grant you many points. Please skip them, unless you have solved everything else.
Don't forget quantifiers!

In all the exercises, d, m are fixed positive integers, with $d \geq 2$.

Exercise 1 *(6 points)*

We consider the map

$$\begin{array}{rccc} f & : & \mathbb{R}^d & \to & \mathbb{R} \\ & & x = (x_1, \dots, x_d) & \to & \left(9 + (x_1^2 - 2)^2\right)\left(1 + \frac{1}{4}||x||_2^2\right). \end{array}$$

1. Compute
   — the first-order critical points of f;
   — its second-order critical points;
   — its global minimizer(s), if it/they exist(s).
2. State and prove a global convergence result for gradient descent over f, with constant stepsize.

Exercise 2 *(8 points + 3 bonus points)*

We want to recover a 1-sparse vector $x_* \in \mathbb{R}^d$, from linear measurements

$$y_* = Ax_* \in \mathbb{R}^m,$$

where $A \in \mathbb{R}^{m\times d}$ is a fixed matrix.

1. Find a property for the matrix A which is a necessary and sufficient condition for this inverse problem to satisfy the uniqueness property.

From now on, we assume that A satisfies this condition.

2. Propose a simple reconstruction algorithm for x_*.

In the rest of the exercise, we study a different reconstruction algorithm (which is widely inferior to the one you just proposed, but is interesting as a simple version of a more sophisticated algorithm, applicable to more difficult problems).
For any $x \in \mathbb{R}^d$, let $i_x \in \{1, \dots, d\}$ be such that

$$|x_{i_x}| = \max_{i=1,\dots,d} |x_i|.$$

(If several values are possible for i_x, we pick one of them, arbitrarily.) We define $H : \mathbb{R}^d \to \mathbb{R}^d$ the map such that, for any $x \in \mathbb{R}^d$,

$$\begin{array}{rcll} H(x)_{i_x} & = & x_{i_x} & \\ H(x)_i & = & 0 & \forall i \in \{1, \dots, d\} \setminus \{i_x\}. \end{array}$$

We also fix some $\tau > 0$.

Now, starting at an arbitrary point $x_0 \in \mathbb{R}^d$, the algorithm consists in defining a sequence of iterates $(x_t)_{t\in\mathbb{N}}$ through the iteration formula

$$x_{t+1} = H\left(x_t - \tau A^T\left(Ax - y_*\right)\right) \tag{Alg}$$

3. This algorithm can be seen as an instance of an algorithm described in the last chapter of the class. Which one ? Provide a short justification of your answer.

For some $m_1, m_2, M > 0$, we assume that

1. the norm of each column of A belongs to $[m_1; m_2]$;
2. the operator norm of A is at most M, i.e.
$$||Az||_2 \leq M||z||_2, \forall z \in \mathbb{R}^d;$$
3. $\tau < \frac{1}{m_2 M}$;
4. $x_* \neq 0$.

4. Let $t \in \mathbb{N}$ be arbitrary.
   a) [Bonus] Show that, for any $i = 1, \dots, d$,
   $$\left|\left((\mathrm{Id} - \tau A^T A)(x_t - x_*)\right)_i\right| \leq (1 + \tau m_2 M)||x_t - x_*||_2.$$
   b) We assume that $x_t \in B\left(x_*, \frac{||x_*||_2}{4}\right)$. Show that
   $$\left|\left(x_t - \tau A^T(Ax_t - y_*)\right)_{i_{x_*}}\right| > \frac{||x_*||_2}{2}$$
   $$\text{and } \forall i \neq i_{x_*}, \quad \left|\left(x_t - \tau A^T(Ax_t - y_*)\right)_i\right| < \frac{||x_*||_2}{2}.$$
   c) Show that x_{t+1} is 1-sparse, with the same non-zero coordinate as x_*.
   d) [Bonus] Show that if $x_t \in B\left(x_*, \frac{||x_*||_2}{8}\right)$, then $x_{t+1} \in B\left(x_*, \frac{||x_*||_2}{4}\right)$.
5. Let $t \in \mathbb{N}$ be arbitrary. We assume that x_t is 1-sparse, with the same non-zero coordinate as x_* : $i_{x_*} = i_{x_t}$.
   a) [Bonus] Show that
   $$\left|\left(x_t - \tau A^T(Ax_t - y_*)\right)_{i_{x_*}} - (x_*)_{i_{x_*}}\right| \leq \left(1 - \tau m_1^2\right)||x_t - x_*||_2.$$
   b) [Bonus] Show that, if $x_t \in B\left(x_*, \frac{||x_*||_2}{4}\right)$, then $||x_{t+1} - x_*||_2 \leq \left(1 - \tau m_1^2\right)||x_t - x_*||_2$.
6. a) Using the previous questions, prove that, if $x_0 \in B\left(x_*, \frac{||x_*||_2}{8}\right)$, then the sequence of iterates $(x_t)_{t\in\mathbb{N}}$ generated by Algorithm (Alg) converges to x_*.
   b) According to the terminology used in class, how would you call the result you just proved ?

Exercise 3 *(8 points + 4.5 bonus points)*

Let $\mathbb{1}_d$ denote the vector in $\mathbb{R}^d$ whose coordinates are all 1, and

$$\mathcal{E}_\perp = \{x \in \mathbb{R}^d, \langle x, \mathbb{1}_d\rangle = 0\}.$$

Let $A \in \mathbb{R}^{m\times d}$ be a matrix such that $A\mathbb{1}_d = 0$.

We want to recover a vector $x_* \in \mathcal{E}_\perp$ from $y_* \overset{def}{=} Ax_*$, knowing that x_* has a small number of "jumps", i.e.

$$\text{Card}\{i \in \{1, \dots, d-1\}, (x_*)_i \neq (x_*)_{i+1}\} \leq r,$$

for some fixed integer $r \ll d$. We consider the convex relaxation

$$\begin{aligned} &\text{minimize } ||x||_{TV} \overset{def}{=} \sum_{i=1}^{d-1} |x_i - x_{i+1}| \\ &\text{over all } x \in \mathcal{E}_\perp \\ &\text{such that } Ax = y_*. \end{aligned} \tag{Min TV}$$

1. Define the word *tightness* in the context of Problem (Min TV).
2. a) Provide a conjecture on the extremal points of

   $$\{x \in \mathcal{E}_\perp, ||x||_{TV} \leq 1\}$$

   which, if true, can intuitively justify that the convex relaxation (Min TV) may be tight.

   b) [Bonus] Prove your conjecture.
3. Show that the dual of (Min TV) is

   $$\begin{aligned} &\text{maximize } \langle y_*, z\rangle \\ &\text{over all } z \in \mathbb{R}^m \\ &\text{such that } \left|\sum_{k=1}^{K} (A^T z)_k\right| \leq 1, \forall K = 1, \dots, d-1. \end{aligned} \tag{Dual TV}$$

   (Note : depending on how exactly you define the dual, you may find a slightly different - but equivalent - dual problem. This is fine ; do not try to come back to the exact expression above.)
4. [Bonus] Find a property ($\mathcal{P}$) over pairs $(x, z) \in \mathcal{E}_\perp \times \mathbb{R}^m$ such that a pair (x, z) is primal-dual optimal for Problems (Min TV) and (Dual TV) if and only if it satisfies ($\mathcal{P}$).

Exercise 4 *(3 points)*

We consider a (real) phase retrieval problem : $v_1, \dots, v_m \in \mathbb{R}^d$ are known vectors, and we want to recover a signal $x_{sol} \in \mathbb{R}^d$ from the absolute value of linear measurements

$$y_k \overset{def}{=} |\langle x_{sol}, v_k\rangle|, k = 1, \dots, m. \tag{1}$$

This problem can be solved using, for instance, the Wirtinger Flow algorithm. We recall that the spectral initialization procedure consists in defining

$$\hat{M} = \frac{1}{m} \sum_{k=1}^{m} y_k^2 v_k v_k^*,$$

and x_0 as (a multiple of) the main eigenvector of $\hat{M}$. This x_0 is an approximation of x_{sol}, with high probability, when $v_1, \dots, v_m \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, I_d)$.
Now, imagine that, in addition to the absolute values in (1), we are also provided with $m' = \frac{d}{2}$ classical linear measurements :

$$z_k \overset{def}{=} \langle x_{sol}, u_k\rangle, k = 1, \dots, m',$$

where $u_1, \dots, u_{m'}$ are also known vectors.
In the setting where $u_1, \dots, u_{m'} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, I_d)$, can you propose an improvement of the spectral procedure, which takes advantage of this additional information ?
(Only a description of the new procedure is expected ; do not try to justify that it provides a higher-quality approximation.)