Non-convex inverse problems

March 1, 2024, 2 hours

You can use any written or printed material.
For each exercise, the number of points is an indication ; it may change. The bonus questions are not directly related to inverse problems. Therefore, they will not earn you many points (and some of them are quite difficult), so you are encouraged to admit the results and skip the questions, unless you have answered all the rest.

Exercise 1

Give an example of a realistic inverse problem which is not in the lecture notes.
(Be precise : I must understand what is the unknown, and what is the information.)
[2 points]

Exercise 2

We consider the problem

$$\begin{aligned} &\text{recover } (x_1, x_2) \in \mathbb{R}^2 \\ &\text{from } y_1 \overset{def}{=} x_1 \\ &\text{and } y_2 \overset{def}{=} \frac{x_2}{1+x_1^2}. \end{aligned}$$

Is reconstruction unique ? Stable ?
[3 points]

Exercise 3

Let $d, k, m \in \mathbb{N}^*$ be fixed, with $k \ll d$.
For any $E \subset \{1, \dots, d\}$, we define $e_E \in \mathbb{R}^d$ the vector such that

$$\begin{aligned} &(e_E)_i = 1, \forall i \in E, \\ &(e_E)_i = -1, \forall i \in \{1, \dots, d\} \setminus E. \end{aligned}$$

We say that a vector $x \in \mathbb{R}^d$ is *k-regular* if there exist $E_1, \dots, E_k \subset \{1, \dots, d\}$ and $s_1, \dots, s_k \in \mathbb{R}$ such that

$$x = s_1 e_{E_1} + \dots + s_k e_{E_k}.$$

Let $\mathcal{E}_k \subset \mathbb{R}^d$ be the set of k-regular vectors.
For given $A \in \mathbb{R}^{m \times d}, y \in \mathbb{R}^m$, we consider the problem

$$\begin{aligned} &\text{recover } x \in \mathbb{R}^d \\ &\text{knowing that } x \in \mathcal{E}_k, \\ &\text{and } Ax = y. \end{aligned} \tag{Regular}$$

1. Is this problem convex or non-convex ?

2. [Bonus] Show that the extremal points of

$$\{x \in \mathbb{R}^d, ||x||_\infty \leq 1\}$$

are the vectors e_E, for all $E \subset \{1, \dots, d\}$.
[We define the infinity norm as usual : for any $x \in \mathbb{R}^d$, $||x||_\infty = \max_{i\leq d} |x_i|$.]

3. Propose a convex relaxation for Problem (Regular).
4. In the context of Problem (Regular), propose a reasonable notion of k-restricted isometry constant for the matrix A.
5. a) Compute the dual problem to your convex relaxation.
   b) Let x_0 be the (unknown) solution of Problem (Regular), that is, a k-regular vector such that $Ax_0 = y$. Let us assume that there exists $c \in \mathbb{R}^m$ such that

$$\begin{aligned} &||A^T c||_1 = 1 \\ \text{and } &(A^T c)_i \geq 0, \forall i \leq d \text{ such that } (x_0)_i = ||x_0||_\infty, \\ &(A^T c)_i \leq 0, \forall i \leq d \text{ such that } (x_0)_i = -||x_0||_\infty, \\ &(A^T c)_i = 0, \forall i \leq d \text{ such that } |(x_0)_i| < ||x_0||_\infty. \end{aligned}$$

   Using the dual certificate strategy, show that your convex relaxation is tight.

[Remark : actually, this relaxation is most often *not* tight.]
[9 points]

Exercise 4

Let $d, m \in \mathbb{N}^*$ be fixed.
We want to identify some vector $x_{sol} \in \mathbb{R}^d \setminus \{0\}$. For this, we assume that we have access to m measurements $y_1, \dots, y_m \in \mathbb{R}$ such that, for all $k \leq m$,

$$y_k = u_k^{(1)} \cos\left(\left\langle v_k^{(1)}, x_{sol} \right\rangle\right) + u_k^{(2)} \cos\left(\left\langle v_k^{(2)}, x_{sol} \right\rangle\right),$$

where $u_k^{(1)}, u_k^{(2)} \in \mathbb{R}, v_k^{(1)}, v_k^{(2)} \in \mathbb{R}^d$ are known.
Since the cosine is an even function, these measurements do not allow to distinguish x_{sol} from $-x_{sol}$, so we are only interested in reconstructing x_{sol} up to sign.

1. a) Propose a smooth unconstrained optimization problem which is equivalent to identifying $\pm x_{sol}$ from $y_1, \dots, y_m$.
   b) Based on this optimization problem, propose an algorithm for reconstructing x_{sol}.

We would like to rigorously analyze the correctness of the algorithm you just proposed, but it is too difficult. As a first step towards an analysis, we replace the objective function from Question 1.a) with

$$\begin{aligned} F : \; \mathbb{R}^d &\to \mathbb{R} \\ x &\to \frac{e^{-2||x||^2}}{2} - e^{-\frac{||x+x_{sol}||^2}{2}} - e^{-\frac{||x-x_{sol}||^2}{2}} + \frac{e^{-2||x_{sol}||^2}}{2}. \end{aligned}$$

[Motivation : F is the expectation of a reasonable objective you could have proposed in 1.a), when the $u_k^{(i)}, v_k^{(i)}$, $k = 1, \dots, m, i = 1, 2$ follow normal distributions.]

We try to recover $\pm x_{sol}$ by running gradient descent on F, starting at a point x_0.[1] This yields a sequence of iterates $(x_t)_{t\in\mathbb{N}}$.
You can use without proof the following formulas : for all $x \in \mathbb{R}^d$,

$$\begin{aligned}
\nabla f(x) &= \left(-2e^{-2||x||^2} + e^{-\frac{||x+x_{sol}||^2}{2}} + e^{-\frac{||x-x_{sol}||^2}{2}}\right) x \\
&\quad + \left(e^{-\frac{||x+x_{sol}||^2}{2}} - e^{-\frac{||x-x_{sol}||^2}{2}}\right) x_{sol}; \\
\mathrm{Hess} f(x) &= \left(-2e^{-2||x||^2} + e^{-\frac{||x+x_{sol}||^2}{2}} + e^{-\frac{||x-x_{sol}||^2}{2}}\right) I_d \\
&\quad + 8e^{-2||x||^2} x x^T \\
&\quad - e^{-\frac{||x+x_{sol}||^2}{2}} (x+x_{sol})(x+x_{sol})^T \\
&\quad - e^{-\frac{||x-x_{sol}||^2}{2}} (x-x_{sol})(x-x_{sol})^T.
\end{aligned}$$

2. a) Compute $\mathrm{Hess} f(x_{sol})$.
   b) Deduce from the previous question that, for all x close to x_{sol},

$$\begin{aligned}
\langle \nabla f(x), x - x_{sol}\rangle = \left(1 - e^{-2||x_{sol}||^2}\right) ||x - x_{sol}||^2 + 4e^{-2||x_{sol}||^2} \langle x_{sol}, x - x_{sol}\rangle^2 \\
+ O(||x - x_{sol}||^3).
\end{aligned}$$

   c) [Bonus] Let $\mu > 0$ be a real number. Show that, for all x close to x_{sol},

$$\begin{aligned}
||x - \mu\nabla f(x) - x_{sol}||^2 \le \left(1 - 2\mu\left(1 - e^{-2||x_{sol}||^2}\right)\right) ||x - x_{sol}||^2 + \mu^2 ||\nabla f(x)||^2 \\
+ O(||x - x_{sol}||^3)
\end{aligned}$$

   d) [Bonus] Show that, for all $\mu > 0$ small enough and all x close enough to x_{sol},

$$||x - \mu\nabla f(x) - x_{sol}||^2 \le \left(1 - \mu\left(1 - e^{-2||x_{sol}||^2}\right)\right) ||x - x_{sol}||^2.$$

   [Hint : you can use without proof the inequality $||\nabla f(x)|| \le 5||x - x_{sol}||$, valid for all $x \in \mathbb{R}^d$.]
   e) Show that there exists $\rho > 0$ such that, for any $x_0 \in B(x_{sol}, \rho)$, $(x_t)_{t\in\mathbb{N}}$ converges to x_{sol} if the stepsize of gradient descent is small enough.
   f) Using the terminology of the lectures, how does one call the result you have proved in the previous question ?
3. a) [Bonus] Show that, for all $\lambda \in \mathbb{R}$,

$$(\nabla f(\lambda x_{sol}) = 0) \iff (\lambda \in \{-1, 0, 1\}).$$

   [Hint : you can use without proof the following property : for all $a, b \in \mathbb{R}$,

$$\left(a e^{-\frac{a^2}{2}} + b e^{-\frac{b^2}{2}} = (a+b) e^{-\frac{(a+b)^2}{2}}\right) \iff (a = 0 \text{ or } b = 0 \text{ or } a = -b). \quad]$$

   b) Compute the first-order critical points of F.
   c) Compute the second-order critical points of F.

[9.5 points]

1. This is an idealized algorithm only : F cannot be computed when x_{sol} is unknown.