Non-convex inverse problems: exercises

Irène Waldspurger

waldspurger@ceremade.dauphine.fr

Initial version: January to March 2023
This version: January to March 2024

Contents

1 Exercises

Exercise 1: linear inverse problems

Let d, m be positive integers, with $d \leq m$. Let $A \in \mathbb{R}^{m\times d}$ be a matrix. For a given $y \in \mathbb{R}^m$, we consider the inverse problem

$$\text{find } x \in \mathbb{R}^d \text{ such that } Ax = y. \tag{Lin-inverse}$$

1. Under which conditions on A and y does Problem (Lin-inverse) have exactly one solution?
2. *(Singular value decomposition)* In this question, we show the existence of orthogonal matrices $U \in \mathbb{R}^{m\times m}, V \in \mathbb{R}^{d\times d}$, and nonnegative numbers $\lambda_1 \geq \dots \geq \lambda_d \in \mathbb{R}^+$, such that

$$A = UDV,$$

with

$$D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & \ddots & \ddots & \\ & & & \lambda_d \\ & & \dots & 0 \\ \vdots & & & \vdots \\ 0 & \dots & \dots & 0 \end{pmatrix}. \tag{1}$$

This decomposition of A is called the *singular value decomposition* (SVD). The numbers $\lambda_1, \dots, \lambda_d$ are the *singular values*. They are uniquely defined.

a) Let $v_1 \in \mathbb{R}^d$ be such that $||v_1||_2 = 1$ and

$$||Av_1||_2 = \max_{v\in\mathbb{R}^d, ||v||_2=1} ||Av||_2.$$

Then, let $v_2, \dots, v_d$ be such that, for any k, $v_k \in \text{Vect}\{v_1, \dots, v_{k-1}\}^\perp$, $||v_k||_2 = 1$, and

$$||Av_k||_2 = \max_{\substack{v\in\text{Vect}\{v_1,\dots,v_{k-1}\}^\perp \\ ||v||_2=1}} ||Av||_2.$$

Show that this definition is valid (i.e. that the maximums exist) and that $(v_1, \dots, v_d)$ is an orthonormal basis of $\mathbb{R}^d$.

b) Show that, for any $k, k' \in \{1, \dots, d\}$ with $k \neq k'$, $\langle Av_k, Av_{k'} \rangle = 0$.
[Hint: assume $k < k'$. Show that, from the definition of v_k, it holds for any $\theta \in \mathbb{R}$ that $||A(\cos(\theta)v_k + \sin(\theta)v_{k'})||_2 \le ||Av_k||_2$. Raise the inequality to the square and show that the derivative of the left-hand side with respect to θ must be 0 at $\theta = 0$.]

c) For any $k = 1, \dots, d$, let us set $\lambda_k = ||Av_k||_2$. Show that the λ_k are nonnegative, and that $\lambda_1 \ge \lambda_2 \ge \dots \ge \lambda_d$.

d) Show that there exists an orthonormal basis $(u_1, \dots, u_m)$ of $\mathbb{R}^m$ such that

$$\forall k \le d, \quad Av_k = \lambda_k u_k.$$

e) Let D be defined as in Equation (1), U be the matrix whose columns are $u_1, \dots, u_m$, and V the matrix whose rows are $v_1, \dots, v_d$. Show that U, V are orthogonal matrices, and

$$A = UDV.$$

f) Show that the singular values are uniquely defined: if $\tilde{U}, \tilde{V}, \tilde{\lambda}_1, \dots, \tilde{\lambda}_d$ is another SVD of A, then $\tilde{\lambda}_k = \lambda_k$ for any k.

3. We assume that A, y satisfy the conditions of Question 1, and denote x_* the solution of Problem (Lin-inverse). For $\epsilon \in \mathbb{R}^m$ such that $y + \epsilon$ also satisfies the conditions of Question 1, we denote x_ϵ the solution of Problem (Lin-inverse) when y is replaced with $y + \epsilon$.

a) Assuming $y \neq 0$, show that, for any ϵ,

$$\frac{||x_\epsilon - x_*||_2}{||x_*||_2} \le \frac{\lambda_1}{\lambda_d} \frac{||\epsilon||_2}{||y||_2}.$$

b) Show that the inequality is tight (that is, it is not true anymore if $\frac{\lambda_1}{\lambda_d}$ is replaced with a smaller constant).

c) Under which condition on λ_1 and λ_d is Problem (Lin-inverse) stable?

Exercise 2: an example of linear inverse problem

Let d be a positive integer, and μ a positive real number.
For a given $y \in \mathbb{R}^d$, we consider the inverse problem

$$\text{find } x \in \mathbb{R}^d,$$
$$\text{such that } x_i + \mu \left(\sum_{k=1}^{d} x_k \right) = y_i, \forall i \in \{1, \dots, d\}.$$

1. Show that, for any y, the problem has exactly one solution.
2. For which values of μ can we say that the problem is stable?

Exercise 3 (*2024 exam*) We consider the problem

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

Is reconstruction unique? Stable?

Exercise 4: intersection of convex sets

Let $d \in \mathbb{N}^*$ be fixed. Let $C_1, \dots, C_S \subset \mathbb{R}^d$ be closed convex non-empty sets. We consider the problem

$$\begin{aligned} &\text{find } x \in \mathbb{R}^d, \\ &\text{such that } x \in C_s, \forall s \leq S. \end{aligned} \tag{2}$$

For any $s \leq S$, we denote P_s the projector onto C_s: for any $z \in \mathbb{R}^d$, $P_s(z)$ is the point of C_s which is at minimal distance from z:

$$||P_s(z) - z||_2 = \min_{a \in C_s} ||a - z||_2.$$

It is a classical result from convex analysis that P_s is well-defined (that is, a point at minimal distance exists, and is unique). We assume that the sets C_s are sufficiently simple so that the corresponding projections can be numerically computed.
The goal of the exercise is to present an algorithm to solve (2).

1. We consider any $s \in \{1, \dots, S\}$.
   a) Show that, for all $z \in \mathbb{R}^d, a \in C_s$,

   $$\langle a - P_s(z), z - P_s(z) \rangle \leq 0$$

   b) Show that, for all $z, z' \in \mathbb{R}^d$,

   $$\langle P_s(z') - P_s(z), z - z' - P_s(z) + P_s(z') \rangle \leq 0$$

c) Show that, for all $z, z' \in \mathbb{R}^d$,

$$||P_s(z) - P_s(z')||_2^2 + ||P_s(z) - P_s(z') - z + z'||_2^2 \leq ||z - z'||_2^2.$$

d) Deduce from the previous question that, for all $z, z' \in \mathbb{R}^d$,

$$||P_s(z) - P_s(z')||_2 \leq ||z - z'||_2,$$

and that the inequality is strict, unless $P_s(z) - P_s(z') = z - z'$.

The algorithm starts with an arbitrary initial point $x_0 \in \mathbb{R}^d$. It then computes iteratively a sequence of iterates $(x_k)_{k\in\mathbb{N}}$ defined by

$$\forall n \in \mathbb{N}, \forall s \in \{1, \dots, S\}, \quad x_{nS+s} = P_s(x_{nS+(s-1)}).$$

We assume that Problem (2) has at least one solution:

$$C_1 \cap C_2 \cap \dots \cap C_S \neq \emptyset.$$

2. a) Show that, for any $x_* \in \cap_{s\leq S} C_s$, the sequence $(||x_k - x_*||_2)_{k\in\mathbb{N}}$ is non-increasing, hence that it converges. Let us call $\ell(x_*) \in \mathbb{R}$ the limit.

   b) Show that $(x_{kS})_{k\in\mathbb{N}}$ has a converging subsequence. We denote $x_\infty \in \mathbb{R}^d$ the limit.

   c) Show that $x_\infty \in \cap_{s\leq S} C_s$.
   [Hint: show that $P_1(x_\infty)$ is a limit point of $(x_{kS+1})_{k\in\mathbb{N}}$, then that, for any $x_* \in \cap_{s\leq S} C_s$,

$$||x_\infty - x_*||_2 = ||P_1(x_\infty) - x_*||_2 = \ell(x_*).$$

   Using Question 1.d), show that $x_\infty \in C_1$. Iterate the reasoning to show that $x_\infty \in C_s$ for any $s \leq S$.]

   d) Show that $x_k \overset{k\to+\infty}{\longrightarrow} x_\infty$.

Exercise 5: real phase retrieval

This exercise is about *real phase retrieval problems*, that is phase retrieval problems where the unknown signal and measurement vectors have *real* (and not *complex*) coordinates.

A real phase retrieval problem is any problem of the form

$$\text{find } x \in \mathbb{R}^d$$

$$\text{such that } |\langle x, v_s\rangle| = y_s, \forall s \leq m, \qquad \text{(Real-PR)}$$

where $v_1, \dots, v_m$ is a known family of vectors of $\mathbb{R}^d$, $y_1, \dots, y_m$ are given and "$|.|$" denotes the absolute value.

Since multiplication by -1 does not change the absolue value, a real phase retrieval problem can, at best, be solved up to multiplication by -1.

We say that a family of vectors $(v_1, \dots, v_m)$ satisfies the *complement property* if, for any $S \subset \{1, \dots, m\}$,

$$\text{Vect}\{v_s\}_{s\in S} = \mathbb{R}^d \quad \text{or} \quad \text{Vect}\{v_s\}_{s\notin S} = \mathbb{R}^d.$$

1. In this question, we show that $(v_1, \dots, v_m)$ satisfies the complement property if and only if, for any $y_1, \dots, y_m$, the solution of Problem (Real-PR) (when it exists) is unique.
   a) Let us assume that $(v_1, \dots, v_m)$ satisfies the complement property. Let $y_1, \dots, y_m$ be any numbers. Let $x, x' \in \mathbb{R}^d$ be such that, for any $s \leq m$,
   $$|\langle x, v_s\rangle| = y_s = |\langle x', v_s\rangle|.$$
   Show that $x = x'$ or $x = -x'$.
   [Hint: apply the complement property for $S = \{s, \langle x, v_s\rangle = \langle x', v_s\rangle\}$.]
   b) Let us assume that $(v_1, \dots, v_m)$ does not satisfy the complement property. Show the existence of $z_1, z_2 \in \mathbb{R}^d \setminus \{0\}$ such that
   $$\forall s \leq m, \quad \langle z_1, v_s\rangle = 0 \quad \text{or} \quad \langle z_2, v_s\rangle = 0.$$
   c) Define $x = z_1 + z_2, x' = z_1 - z_2$ and show that Problem (Real-PR) may have a non-unique solution.
2. a) Show that, if Problem (Real-PR) has a unique solution for any $y_1, \dots, y_m$, then $m \geq 2d - 1$.
   b) Conversely, we assume that $m \geq 2d - 1$. Show that, for almost any $(v_1, \dots, v_m) \in (\mathbb{R}^d)^m$, Problem (Real-PR) has a unique solution for any $y_1, \dots, y_m$.
3. Provide an explicit example of a family $(v_1, v_2, v_3) \in (\mathbb{R}^2)^3$ and of a family $(v_1, v_2, v_3, v_4, v_5) \in (\mathbb{R}^3)^5$ for which Problem (Real-PR) has a unique solution for any $y_1, \dots, y_m$.

2 Answers

Answer of Exercise 1

1. Problem (Lin-inverse) has at least one solution if and only if $y \in \text{Range}(A)$. This solution, which we denote x_*, is unique if the set

$$\{x \in \mathbb{R}^d \text{ such that } Ax = Ax_*\} = \{x_* + h, h \in \text{Ker}(A)\}$$

is the singleton $\{x_*\}$. This happens if and only if A is injective (that is $\text{Ker}(A) = \{0\}$).

2. a) The application $v \in \mathbb{R}^d \to ||Av||_2 \in \mathbb{R}$ is continuous. The unit sphere of $\mathbb{R}^d$ is compact. Therefore, the maximum

$$\max_{v \in \mathbb{R}^d, ||v||_2 = 1} ||Av||_2$$

exists (i.e. there is a vector v_1 at which the maximum is attained). Similarly, for any $k \in \{2, \dots, d\}$, the set

$$\{v \in \text{Vect}\{v_1, \dots, v_{k-1}\}^\perp, |||v||_2 = 1\}$$

is compact (it is a bounded and closed subset of a finite-dimensional vector space), and $v \in \mathbb{R}^d \to ||Av||_2 \in \mathbb{R}$ is still continuous. Therefore, the maximum in the definition of v_k exists.
From the definition, the family $(v_1, \dots, v_d)$ contains d vectors of $\mathbb{R}^d$, which all have unit norm and are orthgonal one to each other: it is an orthonormal basis.

b) Let $k, k' \in \{1, \dots, d\}$ be such that $k \neq k'$. We can assume that $k < k'$. Let us show that

$$\langle Av_k, Av_{k'} \rangle = 0.$$

From the definition of $v_{k'}$,

$$v_{k'} \in \text{Vect}\{v_1, \dots, v_{k'-1}\}^\perp \subset \text{Vect}\{v_k\}^\perp \quad \Rightarrow \quad \langle v_{k'}, v_k \rangle = 0.$$

As a consequence, for any $\theta \in \mathbb{R}$,

$$|| \cos(\theta) v_k + \sin(\theta) v_{k'} ||_2 = \sqrt{\cos^2(\theta) ||v_k||_2^2 + \sin^2(\theta) ||v_{k'}||_2^2} = 1. \quad (3)$$

In addition, v_k is in $\text{Vect}\{v_1, \dots, v_{k-1}\}^\perp$ and $v_{k'}$ is in $\text{Vect}\{v_1, \dots, v_{k'-1}\}^\perp \subset \text{Vect}\{v_1, \dots, v_{k-1}\}^\perp$, so

$$\cos(\theta) v_k + \sin(\theta) v_{k'} \in \text{Vect}\{v_1, \dots, v_{k-1}\}^\perp. \quad (4)$$

Equations (3) and (4), together with the definition of v_k, imply:

$$||A\left(\cos(\theta)v_k + \sin(\theta)v_{k'}\right)||_2 \leq ||Av_k||_2, \quad \forall \theta \in \mathbb{R}.$$

We raise this inequality to the square: for all $\theta \in \mathbb{R}$,

$$\begin{aligned} &||A\left(\cos(\theta)v_k + \sin(\theta)v_{k'}\right)||_2^2 \\ &\quad = \cos^2(\theta)||Av_k||_2^2 + 2\sin(\theta)\cos(\theta)\left\langle Av_k, Av_{k'}\right\rangle + \sin^2(\theta)||Av_{k'}||_2^2 \\ &\quad \leq ||Av_k||_2^2. \end{aligned}$$

This means that the map $\theta \to \cos^2(\theta)||Av_k||_2^2 + 2\sin(\theta)\cos(\theta)\left\langle Av_k, Av_{k'}\right\rangle + \sin^2(\theta)||Av_{k'}||_2^2$ reaches its maximum at $\theta = 0$. In particular, its derivative at 0 must be 0:

$$\begin{aligned} 0 &= -2\cos(0)\sin(0)||Av_k||_2^2 + 2(\cos^2(0) - \sin^2(0))\left\langle Av_k, Av_{k'}\right\rangle \\ &\qquad + 2\sin(0)\cos(0)||Av_{k'}||_2^2 \\ &= 2\left\langle Av_k, Av_{k'}\right\rangle. \end{aligned}$$

Therefore, $\left\langle Av_k, Av_{k'}\right\rangle = 0$.

c) The λ_k are nonnegative because a norm is always nonnegative. To show that $(\lambda_1, \dots, \lambda_d)$ is a nonincreasing sequence, we can reuse a part of the reasoning of the previous question. For any $k, k' \in \{1, \dots, d\}$ with $k < k'$, we have seen that $v_{k'}$ belongs to $\text{Vect}\{v_1, \dots, v_{k-1}\}^\perp$, and $||v_{k'}||_2 = 1$. Hence, from the definition of v_k,

$$\lambda_k = ||Av_k||_2 \geq ||Av_{k'}||_2 = \lambda_{k'}.$$

d) Let D be the smallest index such that $\lambda_D = 0$ (it is possible that $\lambda_k \neq 0$ for all $k \leq d$, in which case we set $D = d+1$).
For any $k = 1, \dots, D-1$, we set

$$u_k = \frac{Av_k}{||Av_k||} = \frac{Av_k}{\lambda_k}.$$

This is an orthonormal family of $\mathbb{R}^m$: for any $k < D$, $||u_k|| = 1$, and for any $k, k' < D$ with $k \neq k'$, it holds

$$\left\langle u_k, u_{k'}\right\rangle = \frac{\left\langle Av_k, Av_{k'}\right\rangle}{\lambda_k \lambda_{k'}} = 0$$

from Question 2.b). We define $u_D, \dots, u_m$ so that $(u_1, \dots, u_m)$ is an orthonormal basis of $\mathbb{R}^m$.
For any $k < D$, we have $Av_k = \lambda_k u_k$ by construction. And for any $k = D, ..., d$, since $\lambda_k = ||Av_k|| = 0$, it also holds $Av_k = 0 = \lambda_k u_k$.

e) The matrices U, V are orthogonal because their columns (resp. rows, for V) form an orthonormal basis of $\mathbb{R}^m$ (resp. $\mathbb{R}^d$).
The equation

$$\forall k \leq d, \quad Av_k = \lambda_k u_k$$

reads, in matricial form,

$$A \begin{pmatrix} v_1 & \dots & v_d \end{pmatrix} = \begin{pmatrix} u_1 & \dots & u_m \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & \ddots & \ddots & \\ & & & \lambda_d \\ & & \dots & 0 \\ \vdots & & & \vdots \\ 0 & \dots & \dots & 0 \end{pmatrix},$$

which is equivalent to

$$AV^T = UD,$$

which is in turn equivalent, since $V^T V = V V^T = \mathrm{Id}$, to

$$A = UDV.$$

f) Let $\tilde{U}, \tilde{V}, \tilde{\lambda}_1, \dots, \tilde{\lambda}_d$ be another SVD of A. Let us denote

$$\tilde{D} = \begin{pmatrix} \tilde{\lambda}_1 & 0 & \dots & 0 \\ 0 & \tilde{\lambda}_2 & & \vdots \\ \vdots & \ddots & \ddots & \\ & & & \tilde{\lambda}_d \\ & & \dots & 0 \\ \vdots & & & \vdots \\ 0 & \dots & \dots & 0 \end{pmatrix}.$$

From the definition of the SVD,

$$\begin{aligned} & A = UDV = \tilde{U}\tilde{D}\tilde{V} \\ \Rightarrow \quad & A^T A = V^T D^T D V = \tilde{V}^T \tilde{D}^T \tilde{D} \tilde{V}. \end{aligned}$$

The matrix $D^T D$ is diagonal, with coefficients on the diagonal $\lambda_1^2, \dots, \lambda_d^2$. The matrices V and V^T are inverse one from each other, since V is an orthogonal matrix. As a consequence, $V^T (D^T D) V$ is the eigenvector decomposition of $A^T A$ and $\lambda_1^2, \dots, \lambda_d^2$ are the eigenvalues of $A^T A$.
For the same reason, $\tilde{\lambda}_1^2, \dots, \tilde{\lambda}_d^2$ are the eigenvalues of $A^T A$. Since the eigenvalues of a matrix are uniquely defined and $\lambda_1^2, \dots, \lambda_d^2$ as well as

$\tilde{\lambda}_1^2, \dots, \tilde{\lambda}_d^2$ are ordered (they are non-increasing sequences), we must have

$$\lambda_1^2 = \tilde{\lambda}_1^2, \quad \dots, \quad \lambda_d^2 = \tilde{\lambda}_d^2,$$

which implies, since the λ_k and $\tilde{\lambda}_k$ are nonnegative,

$$\lambda_1 = \tilde{\lambda}_1, \quad \dots, \quad \lambda_d = \tilde{\lambda}_d,$$

3. a) We assume that A, y and $A, y+\epsilon$ satisfy the conditions of Question 1, that is A is injective, and $y, y+\epsilon$ belong to Range(A).
We consider the SVD of A, as in Question 2. We observe that $\lambda_1 \neq 0, \dots, \lambda_d \neq 0$, otherwise D would not be injective, and A would not be either.
We have

$$\begin{aligned} & UDVx_* = Ax_* = y \quad \text{and} \quad UDVx_\epsilon = Ax_\epsilon = y + \epsilon, \\ \Rightarrow \quad & D(Vx_*) = U^T y \quad \text{and} \quad D(Vx_\epsilon) = U^T(y+\epsilon) = U^T y + U^T \epsilon. \end{aligned} \tag{5}$$

We respectively denote $(x_{V,k})_{k \leq d}$, $(x_{V,k}^{(\epsilon)})_{k \leq d}$, $(y_{U,k})_{k \leq m}$ and $(\epsilon_{U,k})_{k \leq m}$ the coordinates of Vx_*, Vx_ϵ, $U^T y$ and $U^T \epsilon$. From Equation (5), for all $k \leq d$,

$$\begin{aligned} & \lambda_k x_{V,k} = y_{U,k} \quad \text{and} \quad \lambda_k x_{V,k}^{(\epsilon)} = y_{U,k} + \epsilon_{U,k}, \\ \Rightarrow \quad & x_{V,k} = \frac{y_{U,k}}{\lambda_k} \quad \text{and} \quad x_{V,k}^{(\epsilon)} = \frac{y_{U,k}}{\lambda_k} + \frac{\epsilon_{U,k}}{\lambda_k} \end{aligned}$$

and, for all $k = d+1, \dots, m$,

$$y_{U,k} = \epsilon_{U,k} = 0.$$

From these equalities we deduce

$$\begin{aligned} ||Vx_*||_2 &= \left(\sum_{k=1}^d x_{V,k}^2 \right)^{1/2} = \left(\sum_{k=1}^d \frac{y_{U,k}^2}{\lambda_k^2} \right)^{1/2} \\ &\geq \left(\sum_{k=1}^d \frac{y_{U,k}^2}{\lambda_1^2} \right)^{1/2} = \frac{1}{\lambda_1} \left(\sum_{k=1}^m y_{U,k}^2 \right)^{1/2} = \frac{||U^T y||_2}{\lambda_1} \end{aligned}$$

and

$$||V(x_* - x_\epsilon)||_2 = \left(\sum_{k=1}^{d} \left(x_{V,k} - x_{V,k}^{(\epsilon)} \right)^2 \right)^{1/2}$$

$$= \left(\sum_{k=1}^{d} \frac{\epsilon_{U,k}^2}{\lambda_k^2} \right)^{1/2} \leq \left(\sum_{k=1}^{d} \frac{\epsilon_{U,k}^2}{\lambda_d^2} \right)^{1/2}$$

$$= \frac{1}{\lambda_d} \left(\sum_{k=1}^{m} \epsilon_{U,k}^2 \right)^{1/2} = \frac{||U^T \epsilon||_2}{\lambda_d}.$$

Therefore,

$$\frac{||V(x_* - x_\epsilon)||_2}{||V x_*||_2} \leq \frac{\lambda_1}{\lambda_d} \frac{||U^T \epsilon||_2}{||U^T y||_2}$$

and, since V, U are orthogonal matrices, hence preserve the norm of vectors,

$$\frac{||x_* - x_\epsilon||_2}{||x_*||_2} \leq \frac{\lambda_1}{\lambda_d} \frac{||\epsilon||_2}{||y||_2}.$$

b) Let us consider the following y and ϵ:

$$y = U e_1, \quad \epsilon = U e_d,$$

where e_1, e_d respectively denote the first and d-th vector in the canonical basis of $\mathbb{R}^m$. Then

$$x_* = \frac{1}{\lambda_1} V^T \tilde{e}_1, \quad x_\epsilon = \frac{1}{\lambda_1} V^T \tilde{e}_1 + \frac{1}{\lambda_d} V^T \tilde{e}_d,$$

where $\tilde{e}_1, \tilde{e}_d$ respectively denote the first and d-th vector in the canonical basis of $\mathbb{R}^d$. Therefore,

$$\frac{||x_* - x_\epsilon||_2}{||x_*||_2} = \frac{\lambda_1}{\lambda_d} \frac{||V^T \tilde{e}_d||_2}{||V^T \tilde{e}_1||_2} = \frac{\lambda_1}{\lambda_d} = \frac{\lambda_1}{\lambda_d} \frac{||\epsilon||_2}{||y||_2}.$$

c) The inverse problem is stable if $\frac{\lambda_1}{\lambda_d}$ is of order 1 (say ≤ 10).

Answer of Exercise 2

1. We are exactly in the same setting as the previous exercise, with

$$A = \begin{pmatrix} 1 & \mu & \dots & \mu \\ \mu & 1 & \dots & \\ \vdots & & \ddots & \vdots \\ \mu & & \dots & 1 \end{pmatrix}.$$

   According to Question 1 of the previous exercise, we must show that A is injective and surjective. Given that A is square, it is enough to show that A is injective.
   To show that A is injective, we consider $z \in \mathrm{Ker}(A)$ and show that, necessarily, $z = 0$. From the definition of the kernel,

$$z_i + \mu \left(\sum_{k=1}^{d} z_k \right) = 0, \forall i \in \{1, \dots, d\}.$$

   Therefore, all coordinates of z are equal:

$$z_1 = z_2 = \dots = z_d = -\mu \left(\sum_{k=1}^{d} z_k \right).$$

   We plug this into the first equation:

$$(1 + d\mu) z_1 = 0.$$

   Since $\mu > 0$, we must have $z_1 = 0$, and therefore $z_2 = \dots = z_d = 0$, that is $z = 0$.

2. Following the previous exercise, we compute the singular value decomposition of A. As A is a symmetric matrix, its singular values are the absolute values of its eigenvalues. Let us compute the eigenvalues.
   Let for the moment $\lambda \in \mathbb{R}$ be any eigenvalue, and let z be an associated eigenvector. From the definition of A,

$$z_i + \mu \left(\sum_{k=1}^{d} z_k \right) = \lambda z_i, \forall i \in \{1, \dots, d\}.$$

Therefore, if $\lambda \neq 1$, it holds

$$z_1 = z_2 = \cdots = z_d = \frac{\mu}{\lambda - 1}\left(\sum_{k=1}^{d} z_k\right),$$

which means that z is a constant vector.
Conversely, if z is a constant vector, we see that it is an eigenvector, with eigenvalue $1 + d\mu$.
Since the set of constant vectors has dimension 1, we conclude that there is exactly one eigenvalue different from 1, which is $1 + d\mu$ and has multiplicity 1.
Since A are d eigenvalues (when counted with multiplicity), the only other eigenvalue is 1, with multiplicity $d - 1$.
The eigenvalues are nonnegative, so they are the same as the singular values.
From the previous exercise, the inverse problem is stable if the ratio between that largest and smallest singular values is of order 1, that is if $1 + d\mu$ is of order 1. In other words, reconstruction is stable when μ is at most of order $\frac{1}{d}$.

Answer of Exercise 3

[Caution: this is a *non-linear* inverse problem. Therefore, it cannot be analyzed using the results on *linear* inverse problems.]

Reconstruction is unique: for any $(x_1, x_2) \in \mathbb{R}^2$ and associated measurements (y_1, y_2), it holds $(x_1, x_2) = (y_1, (1 + y_1^2)y_2)$. Therefore, the measurements (y_1, y_2) uniquely determine (x_1, x_2).
Reconstruction is not stable. Indeed, for any $\epsilon > 0$, there exist pairs (x_1, x_2) and (x'_1, x'_2), with associated measurements $(y_1, y_2), (y'_1, y'_2)$ such that

$$\frac{||(y_1, y_2) - (y'_1, y'_2)||_2}{||(y_1, y_2)||_2} \leq \epsilon \frac{||(x_1, x_2) - (x'_1, x'_2)||_2}{||(x_1, x_2)||_2}.$$

To show it, we can consider the pairs $(x_1, x_2) = (t, t)$ and $(x'_1, x'_2) = (t, 0)$, for some $t > 0$ to be defined later. Then

$$\frac{||(x_1, x_2) - (x'_1, x'_2)||_2}{||(x_1, x_2)||_2} = \frac{t}{\sqrt{2}t} = \frac{1}{\sqrt{2}}$$

while

$$
\begin{aligned}
\frac{||(y_1,y_2)-(y_1',y_2')||_2}{||(y_1,y_2)||_2} &= \frac{\frac{t}{1+t^2}}{t\sqrt{1+\frac{1}{(1+t^2)^2}}} \\
&= \frac{1}{(1+t^2)\sqrt{1+\frac{1}{(1+t^2)^2}}} \\
&\leq \frac{1}{t^2}.
\end{aligned}
$$

Consequently, if $t \geq \frac{2^{1/4}}{\epsilon^{1/2}}$, it holds

$$
\frac{||(y_1,y_2)-(y_1',y_2')||_2}{||(y_1,y_2)||_2} \leq \epsilon \frac{||(x_1,x_2)-(x_1',x_2')||_2}{||(x_1,x_2)||_2}.
$$

Answer of Exercise 4

1. a) Let $z \in \mathbb{R}^d, a \in C_s$ be fixed. For any $\epsilon \in [0;1]$, the vector

$$
(1-\epsilon)P_s(z) + \epsilon a
$$

belongs to C_s, since $P_s(z)$ and a belong to C_s and C_s is convex. Therefore, from the definition of the projection,

$$
\begin{aligned}
||P_s(z)-z||_2^2 &\leq ||(1-\epsilon)P_s(z)+\epsilon a - z||^2 \\
&= ||P_s(z)-z||^2 - \epsilon \langle a - P_s(z), z - P_s(z)\rangle \\
&\quad + \epsilon^2 ||a - P_s(z)||_2^2.
\end{aligned}
$$

Therefore, for any $\epsilon \in]0;1]$,

$$
\langle a - P_s(z), z - P_s(z)\rangle \leq \epsilon ||a - P_s(z)||_2^2.
$$

If we let ϵ go to 0, we get that $\langle a - P_s(z), z - P_s(z)\rangle \leq 0$.

b) Let $z, z' \in \mathbb{R}^d$ be fixed. We apply the previous question to $a = P_s(z')$, then to $a = P_s(z)$:

$$
\begin{aligned}
\langle P_s(z') - P_s(z), z - P_s(z)\rangle \leq 0, \\
\langle P_s(z) - P_s(z'), z' - P_s(z')\rangle \leq 0.
\end{aligned}
$$

We sum the two inequalities and get the desired result.

c) Let $z, z' \in \mathbb{R}^d$ be fixed.

$$\begin{aligned}
||P_s(z) - P_s(z')||_2^2 &+ ||P_s(z) - P_s(z') - z + z'||_2^2 \\
&= ||z - z'||_2^2 + 2||P_s(z) - P_s(z')||_2^2 + 2\left\langle P_s(z) - P_s(z'), z' - z \right\rangle \\
&= ||z - z'||_2^2 + 2\left\langle P_s(z') - P_s(z), z - z' - P_s(z) + P_s(z') \right\rangle \\
&\leq ||z - z'||_2^2.
\end{aligned}$$

The last inequality is a consequence of Question 1.b).

d) For all $z, z' \in \mathbb{R}^d$, from the previous question, since $||P_s(z) - P_s(z') - z + z'||_2^2 \geq 0$, we must have

$$||P_s(z) - P_s(z')||_2^2 \leq ||z - z'||_2^2,$$

hence $||P_s(z) - P_s(z')||_2 \leq ||z - z'||_2$. In addition, if the inequality is not strict, it must hold $||P_s(z) - P_s(z') - z + z'||_2^2 = 0$, so that $P_s(z) - P_s(z') - z + z' = 0$, hence $P_s(z) - P_s(z') = z - z'$.

2. a) Let $k \in \mathbb{N}$ be fixed. Let $n \in \mathbb{N}, s \in \{1, \dots, S\}$ be such that $k = nS + (s-1)$. Then $x_{k+1} = x_{nS+s} = P_s(x_k)$. In addition, $x_* = P_s(x_*)$ because x_* is in C_s, so

$$||x_{k+1} - x_*||_2 = ||P_s(x_k) - P_s(x_*)||_2 \leq ||x_k - x_*||_2.$$

The last inequality is true from Question 1.d). The sequence is therefore nonincreasing. It has a limit as any nonnegative nonincreasing sequence of real numbers has a limit.

b) The sequence $(x_{kS})_{k\in\mathbb{N}}$ is bounded: for any k,

$$\begin{aligned}
||x_{kS}||_2 &\leq ||x_*||_2 + ||x_{kS} - x_*||_2 \\
&\leq ||x_*||_2 + ||x_0 - x_*||_2.
\end{aligned}$$

From Bolzano-Weierstrass theorem, it has a converging subsequence.

c) As P_1 is continuous (from Question 1.d), it is even 1-Lipschitz) and x_∞ is a limit point of $(x_{kS})_{k\in\mathbb{N}}$, $P_1(x_\infty)$ is a limit point of $(P_1(x_{kS}))_{k\in\mathbb{N}} = (x_{kS+1})_{k\in\mathbb{N}}$.

Since $||x_{kS} - x_*||_2 \overset{k\to+\infty}{\longrightarrow} \ell(x_*)$ and $||x_{kS+1} - x_*||_2 \overset{k\to+\infty}{\longrightarrow} \ell(x_*)$, we must have

$$||x_\infty - x_*||_2 = \ell(x_*),$$

$$||P_1(x_\infty) - P_1(x_*)||_2 = ||P_1(x_\infty) - x_*||_2 = \ell(x_*).$$

This implies that $||x_\infty - x_*||_2 = ||P_1(x_\infty) - P_1(x_*)||_2$. From Question 1.d), we must have

$$x_\infty - x_* = P_1(x_\infty) - P_1(x_*) = P_1(x_\infty) - x_*,$$

so that $x_\infty = P_1(x_\infty)$, which is equivalent to $x_\infty \in C_1$.
We can reapply iteratively the same reasoning: as $x_\infty = P_1(x_\infty)$ is a limit point of $(x_{kS+1})_{k\in\mathbb{N}}$, $P_2(x_\infty)$ is a limit point of $(x_{kS+2})_{k\in\mathbb{N}}$, which allows to show that $||x_\infty - x_*||_2 = ||P_2(x_\infty) - P_2(x_*)||_2$, hence $x_\infty = P_2(x_\infty)$, so that $x_\infty \in C_2$. And so on.

d) As x_∞ belongs to $\cap_{s\leq S} C_s$, Question 2.a) tells us that $(||x_k - x_\infty||_2)_{k\in\mathbb{N}}$ is nonincreasing. This sequence has a subsequence which goes to 0 (since x_∞ is a limit point of $(x_k)_{k\in\mathbb{N}}$). Therefore, the whole sequence goes to 0, which means that $x_k \overset{k\to+\infty}{\longrightarrow} x_\infty$.