

Non-convex inverse problems: exercises

Irène Waldspurger

waldspurger@ceremade.dauphine.fr

January to February 2025

Contents

1 Exercises	3
Exercise 1: linear inverse problems	3
Exercise 2: an example of linear inverse problem	4
Exercise 3	5
Exercise 4: intersection of convex sets	5
Exercise 5: real phase retrieval	6
Exercise 6: correctness guarantees for Basis Pursuit	8
Exercise 7: guarantees for nuclear norm minimization	9
Exercise 8: Prony's method for super-resolution	12
Exercise 9: super-resolution as a semidefinite program	13
Exercise 10	15
Exercise 11: alternating projections for phase retrieval	17
Exercise 12: rank 1 approximation (local convergence)	18
Exercise 13: rank 1 approximation (global convergence)	19
2 Answers	21
Answer of Exercise 1	21
Answer of Exercise 2	26
Answer of Exercise 3	27
Answer of Exercise 4	28
Answer of Exercise 5	30
Answer of Exercise 6	31

Answer of Exercise 7	34
Answer of Exercise 9	37
Answer of Exercise 10	41
Answer of Exercise 11	44
Answer of Exercise 12	45
Answer of Exercise 13	51

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

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

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

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

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 $\|A(\cos(\theta)v_k + \sin(\theta)v_{k'})\|_2 \leq \|Av_k\|_2$. 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 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d$.
- d) Show that there exists an orthonormal basis (u_1, \dots, u_m) of \mathbb{R}^m such that

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

- 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 ,

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

- 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

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

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

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

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 $x_k \xrightarrow{k \rightarrow +\infty} x_\infty$.

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*, as in class) coordinates.

A real phase retrieval problem is any problem of the form

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

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

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

Since multiplication by -1 does not change the absolute 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 (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 .

Exercise 6: correctness guarantees for Basis Pursuit

Let d, m, k be positive integers. For some matrix $A \in \mathbb{R}^{m \times d}$, we consider the problem

$$\begin{aligned} & \text{minimize } \|x\|_1 \\ & \text{for } x \in \mathbb{R}^d \\ & \text{such that } Ax = y. \end{aligned} \quad (\text{Basis Pursuit})$$

We assume that the $4k$ -restricted isometry constant of A satisfies

$$\delta_{4k} < \frac{1}{4}.$$

Let x_* be any vector with at most k non-zero coordinates. We consider Problem (Basis Pursuit) for $y = Ax_*$. Let x_{BP} be any solution. The goal of the exercise is to show that, necessarily,

$$x_{BP} = x_*.$$

1. We define

$$\begin{aligned} h &= x_{BP} - x_*, \\ T_* &= \{i, x_{*i} \neq 0\}. \end{aligned}$$

Show that

$$\|h_{T_*^c}\|_1 \leq \|h_{T_*}\|_1.$$

(For any vector $z \in \mathbb{R}^d$ and $E \subset \{1, \dots, d\}$, z_E is the vector obtained from z by setting to 0 all coordinates corresponding to indices outside E .)

2. Up to permuting the coordinates of x_*, x_{BP} and the columns of A , we can assume that

$$T_* = \{1, 2, \dots, \text{Card}(T_*)\}$$

and the coordinates of h are non-increasing, in absolute value, outside T_* :

$$|h_{\text{Card}(T_*)+1}| \geq |h_{\text{Card}(T_*)+2}| \geq \dots \geq |h_d|.$$

Let us partition $\{\text{Card}(T_*) + 1, \dots, d\}$ into sets T_1, T_2, \dots, T_L of size $3k$:

$$T_1 = \{\text{Card}(T_*) + 1, \dots, \text{Card}(T_*) + 3k\},$$

$$T_2 = \{\text{Card}(T_*) + 3k + 1, \dots, \text{Card}(T_*) + 6k\},$$

...

$$T_L = \{\text{Card}(T_*) + 3(L-1)k + 1, \dots, d\}.$$

a) Show that, for any $l \in \{2, \dots, L\}$,

$$\|h_{T_l}\|_2^2 \leq \frac{\|h_{T_{l-1}}\|_1^2}{3k}.$$

[Hint: for each $s \in T_l$, show that $|h_s| \leq \frac{\|h_{T_{l-1}}\|_1}{3k}$.]

b) Show that

$$\sum_{l=2}^L \|h_{T_l}\|_2 \leq \frac{\|h_{T_*}\|_1}{\sqrt{3k}}.$$

c) Deduce from the last question that

$$\sum_{l=2}^L \|h_{T_l}\|_2 \leq \frac{\|h_{T_*}\|_2}{\sqrt{3}}.$$

3. a) Show that $Ah = 0$.

b) Show that

$$\|Ah\|_2 \geq (1 - \delta_{4k})\|h_{T_* \cup T_1}\|_2 - (1 + \delta_{4k}) \sum_{l=2}^L \|h_{T_l}\|_2.$$

c) Conclude.

Exercise 7: guarantees for nuclear norm minimization

Let d_1, d_2, m, r be positive integers. For some linear operator $\mathcal{L} : \mathbb{R}^{d_1 \times d_2} \rightarrow \mathbb{R}^m$, we consider the problem

$$\begin{aligned} & \text{minimize } \|X\|_* \\ & \text{for } X \in \mathbb{R}^{d_1 \times d_2} \\ & \text{such that } \mathcal{L}(X) = y. \end{aligned} \quad (\text{Nuclear-min})$$

We assume that the $5r$ -restricted isometry constant of \mathcal{L} satisfies

$$\delta_{5r} < \frac{1}{10}.$$

Let X_* be a matrix with rank at most r . Let X_{NM} be a solution of Problem (Nuclear-min) with $y = \mathcal{L}(X_*)$. The goal of the exercise is to show that

$$X_{NM} = X_*.$$

To simplify notation, we assume $d_1 \geq d_2$. If we multiply the matrices to the left and to the right by suitably chosen orthogonal matrices (the inverse of the orthogonal matrices of the SVD of X_*), we can assume that X_* is diagonal:

$$X_* = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & \ddots & \ddots & \\ & & & \lambda_{d_2} \\ & & & 0 \\ \vdots & & & \vdots \\ 0 & \dots & \dots & 0 \end{pmatrix},$$

with the λ_s are nonnegative and ordered: $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{d_2} \geq 0$.

1. Show that $\lambda_{r+1} = \dots = \lambda_{d_2} = 0$.

We set $H = X_{NM} - X_*$ and write its block decomposition

$$H = \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} \begin{matrix} \xleftrightarrow{r} & \xleftrightarrow{d_2-r} \\ \updownarrow r & \updownarrow d_1-r \end{matrix}$$

We set

$$H_0 = \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & 0 \end{pmatrix} \quad \text{and} \quad H_c = \begin{pmatrix} 0 & 0 \\ 0 & H_{22} \end{pmatrix}.$$

We assume that H_{22} is diagonal, with nonnegative ordered diagonal entries. (This is only for simplicity. In the general case, the same reasoning is valid; it suffices to add at the right place multiplications by the orthogonal matrices appearing in the SVD of H_{22} .)

$$H_{22} = \begin{pmatrix} \mu_1 & 0 & \dots & 0 \\ 0 & \mu_2 & & \vdots \\ \vdots & \ddots & \ddots & \\ & & & \mu_{d_2-r} \\ & & & 0 \\ \vdots & & & \vdots \\ 0 & \dots & \dots & 0 \end{pmatrix}, \quad \text{with } \mu_1 \geq \dots \geq \mu_{d_2-r} \geq 0.$$

We define matrices $H_{c,1}, \dots, H_{c,L}$ such that, for any l , $H_{c,l}$ is equal to H_c , except that coefficients μ_s have been replaced with 0 for all

$$s \notin \{3(l-1)r+1, \dots, 3lr\}.$$

With this definition, $H_{c,1}, \dots, H_{c,L}$ are a sequence of diagonal matrices, such that

$$H_c = \sum_{l=1}^L H_{c,l}.$$

2. Show that

$$\|H_0\|_* \geq \|H_c\|_*.$$

[Hint: $\|X_* + H_c\|_* = \|X_*\|_* + \|H_c\|_*$.]

3. a) Following the reasoning of the previous exercise, show that

$$\sum_{l=2}^L \|H_{c,l}\|_F \leq \frac{\|H_0\|_*}{\sqrt{3r}}.$$

b) Show that $\text{rank}(H_0) \leq 2r$ and

$$\|H_0\|_* \leq \sqrt{2r} \|H_0\|_F.$$

c) Deduce that

$$\sum_{l=2}^L \|H_{c,l}\|_F \leq \sqrt{\frac{2}{3}} \|H_0\|_F.$$

4. a) Show that

$$\|\mathcal{L}(H)\|_2 \geq (1 - \delta_{5r}) \|H_0 + H_{c,1}\|_F - (1 + \delta_{5r}) \sum_{l=2}^L \|H_{c,l}\|_F.$$

b) Conclude.

Exercise 8: Prony's method for super-resolution

Let $S \in \mathbb{N}^*$ be fixed. We want to recover a measure

$$\mu_0 = \sum_{s=1}^S a_s \delta_{\tau_s},$$

where a_1, \dots, a_S are non-zero complex numbers, and τ_1, \dots, τ_S are distinct elements of $[0; 1[$. We assume that we have access to its $2S$ lowest-frequency Fourier coefficients:

$$\hat{\mu}_0[k] = \int_0^1 e^{-2\pi i k t} d\mu(t) = \sum_{s=1}^S a_s e^{-2\pi i k \tau_s}, \quad \text{for } k = -(S-1), \dots, S.$$

In this exercise, we present a purely non-convex algorithm to perform the reconstruction, called Prony's method.

1. Show that there exists a unique polynomial P with degree S and leading coefficient equal to 1 such that

$$P(e^{2\pi i \tau_s}) = 0, \quad \forall s = 1, \dots, S$$

Express it as a function of τ_1, \dots, τ_S .

2. Let P be the polynomial defined in the previous question. We call $p_0, \dots, p_S \in \mathbb{C}$ its coefficients:

$$P(X) = \sum_{s=0}^S p_s X^s.$$

The goal is to show that $p \stackrel{\text{def}}{=} \begin{pmatrix} p_0 \\ \vdots \\ p_S \end{pmatrix}$ is the unique (up to scalar multiplication) element in the kernel of

$$M \stackrel{\text{def}}{=} \begin{pmatrix} \overline{\hat{\mu}_0[-(S-1)]} & \overline{\hat{\mu}_0[-(S-2)]} & \dots & \overline{\hat{\mu}_0[1]} \\ \overline{\hat{\mu}_0[-(S-2)]} & \overline{\hat{\mu}_0[-(S-3)]} & \dots & \overline{\hat{\mu}_0[2]} \\ \vdots & & & \vdots \\ \overline{\hat{\mu}_0[0]} & \overline{\hat{\mu}_0[1]} & \dots & \overline{\hat{\mu}_0[S]} \end{pmatrix}.$$

- a) Show that $p \in \text{Ker}(M)$.

b) We now prove uniqueness. Let $q = \begin{pmatrix} q_0 \\ \vdots \\ q_S \end{pmatrix}$ be in $\text{Ker}(M)$. We define

$$Q(X) = \sum_{s=0}^S q_s X^s.$$

Show that, for any $d = 0, \dots, S-1$,

$$\sum_{s=1}^S e^{-2\pi i d \tau_s} \overline{a_s} Q(e^{2\pi i \tau_s}) = 0.$$

c) Deduce from the previous question that $q = \lambda p$ for some $\lambda \in \mathbb{C}$.
[Hint: use the fact that the so-called *Vandermonde matrix*

$$\begin{pmatrix} 1 & 1 & \dots & 1 \\ e^{-2\pi i \tau_1} & e^{-2\pi i \tau_2} & \dots & e^{-2\pi i \tau_S} \\ \vdots & & & \vdots \\ e^{-2\pi i (S-1) \tau_1} & e^{-2\pi i (S-1) \tau_2} & \dots & e^{-2\pi i (S-1) \tau_S} \end{pmatrix}$$

is invertible.]

3. Using the previous question, propose an algorithm to recover μ_0 .

Compared to the total variation approach seen in class, this algorithm is much simpler. In addition, it succeeds whatever the values of $a_1, \dots, a_S, \tau_1, \dots, \tau_S$. However, it is difficult to use as such in practice, since it is very sensitive to noise, and therefore requires a high precision on the measures $\hat{\mu}_0[k]$. In addition, it cannot handle some natural generalizations of the problem, like the case where some Fourier measurements are missing.

Exercise 9: super-resolution as a semidefinite program

In this exercise, we discuss one method for solving the total variation minimization problem

$$\begin{aligned} & \text{minimize } \|\mu\|_{TV} \\ & \text{for } \mu \in \mathcal{M}([0; 1]), \\ & \text{such that } \hat{\mu}[k] = y_k, \forall k = -N, \dots, N. \end{aligned} \tag{Min TV}$$

In the lecture, we have introduced the dual of (Min TV):

$$\text{maximize } \text{Re} \langle z, y \rangle$$

$$\begin{aligned} & \text{for } z \in \mathbb{C}^{2N+1} && (\text{Dual TV}) \\ & \text{such that } \left| \sum_{k=-N}^N z_k e^{2\pi i k t} \right| \leq 1, \forall t \in \mathbb{R}. \end{aligned}$$

We have said that both problems have the same optimal value, and (more or less) that the minimizers of **(Min TV)** can be recovered from the maximizers of **(Dual TV)**. We can therefore focus on solving **(Dual TV)**, which is a convex problem with an infinite number of constraints.

We admit the following result.

Theorem 1 : Fejér-Riesz

Let $P(e^{2\pi i t}) = \sum_{k=-2N}^{2N} p_k e^{2\pi i k t}$ be a trigonometric polynomial with degree at most $2N$. The following two properties are equivalent.

1. P has real nonnegative values on the unit circle (that is, $P(e^{2\pi i t}) \in \mathbb{R}^+$ for all $t \in \mathbb{R}$).
2. There exists a finite number of trigonometric polynomials Q_1, \dots, Q_n , each with degree at most N , such that $P(e^{2\pi i t}) = \sum_{k=1}^n |Q_k(e^{2\pi i t})|^2$.

1. Let $z \in \mathbb{C}^{2N+1}$ be any vector. Show that $\left| \sum_{k=-N}^N z_k e^{2\pi i k t} \right| \leq 1$ for all $t \in \mathbb{R}$ if and only if there exists a finite number of trigonometric polynomials P_1, \dots, P_n with degree at most N such that

$$\left| \sum_{k=-N}^N z_k e^{2\pi i k t} \right|^2 + \sum_{l=1}^n |P_l(e^{2\pi i t})|^2 = 1, \quad \forall t \in \mathbb{R}. \quad (3)$$

2. Let P_1, \dots, P_n be trigonometric polynomials with degree at most N . Let $p^{(1)}, \dots, p^{(n)} \in \mathbb{C}^{2N+1}$ be the vectors of their coefficients:

$$P_l(e^{2\pi i t}) = \sum_{k=-N}^N p_k^{(l)} e^{2\pi i k t}.$$

Show that the polynomials satisfy Equality (3) if and only if the matrix $A = zz^* + \sum_{l=1}^n p^{(l)} p^{(l)*} \in \mathbb{C}^{(2N+1) \times (2N+1)}$ satisfies, for all $d =$

$$-2N, \dots, 2N,$$

$$\sum_{k=1+\max(0,d)}^{2N+1-\max(0,-d)} A_{k,k-d} = 0 \text{ if } d \neq 0$$

$$= 1 \text{ if } d = 0.$$

3. Show that a matrix $A \in \mathbb{C}^{(2N+1) \times (2N+1)}$ can be written as $A = zz^* + \sum_{l=1}^n p^{(l)} p^{(l)*}$ for some vectors $p^{(1)}, \dots, p^{(n)} \in \mathbb{C}^{2N+1}$ if and only if $A - zz^* \succeq 0$.
4. Show that a matrix $A \in \mathbb{C}^{(2N+1) \times (2N+1)}$ satisfies the inequality $A - zz^* \succeq 0$ if and only if

$$\left(\begin{array}{c|c} A & \begin{matrix} z_{-N} \\ \vdots \\ z_N \end{matrix} \\ \hline \begin{matrix} \bar{z}_{-N} & \dots & \bar{z}_N \end{matrix} & 1 \end{array} \right) \succeq 0.$$

5. Deduce from the previous questions that Problem (Dual TV) is equivalent to

$$\begin{aligned} & \text{maximize } \operatorname{Re} \langle z, y \rangle \\ & \text{over all } z \in \mathbb{C}^{2N+1}, A \in \mathbb{C}^{(2N+1) \times (2N+1)} \\ & \text{such that } \sum_{k=1+\max(0,d)}^{2N+1-\max(0,-d)} A_{k,k-d} = 0 \text{ for all } d \in \{-2N, \dots, 2N\} \setminus \{0\}, \\ & \sum_{k=1}^{2N+1} A_{k,k} = 1, \\ & \text{and } \left(\begin{array}{c|c} A & z \\ \hline z^* & 1 \end{array} \right) \succeq 0, \end{aligned}$$

which is a classical (finite-dimensional) semidefinite optimization problem.

Exercise 10 (2024 exam)

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. Show that the extremal points of $\{x \in \mathbb{R}^d, \|x\|_\infty \leq 1\}$ are the vectors e_E , for all $E \subset \{1, \dots, d\}$.
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) In this question, we assume $k = 1$ and $y \neq 0$. Let x_0 be the (unknown) solution of Problem (Regular) (that is, a 1-regular vector such that $Ax_0 = y$). Let us assume that there exists $c \in \mathbb{R}^m$ such that

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

Using the dual certificate strategy, show that x_0 is the unique minimizer of your convex relaxation.

[Remark: actually, for $k > 1$, and even for many ground truths when $k = 1$, this relaxation is most often *not* tight. The goal of the exercise is to get more familiarity with the notions of convex relaxations and dual certificates; the considered convex relaxation is not interesting per se.]

Exercise 11: alternating projections for phase retrieval

We consider a generic phase retrieval problem:

$$\begin{aligned} &\text{find } x \in \mathbb{C}^d \\ &\text{such that } |L_s(x)| = y_s, \forall s \leq m, \end{aligned} \quad (\text{PR})$$

where $L_1, \dots, L_m : \mathbb{C}^d \rightarrow \mathbb{C}$ are known linear maps, and $y_1, \dots, y_m \in \mathbb{R}^+$ are given.

We define $\mathcal{A} : x \in \mathbb{C}^d \rightarrow (L_s(x))_{s=1, \dots, m} \in \mathbb{C}^m$ and

$$\mathcal{E} = \{h \in \mathbb{C}^m \text{ such that } |h_s| = y_s, \forall s = 1, \dots, m\}.$$

1. a) Show that, if x is a solution of (PR), then $\mathcal{A}(x)$ is a solution of the following problem:

$$\text{find } z \in \text{Range}(\mathcal{A}) \cap \mathcal{E}. \quad (\text{Set intersection})$$

- b) Conversely, show that, if z is a solution of (Set intersection), then $z = \mathcal{A}(x)$ for some solution x of (PR).

This shows that, to solve Problem (PR), it suffices to solve (Set intersection).

2. Give the explicit expression of a function $\text{proj}_{\mathcal{E}} : \mathbb{C}^m \rightarrow \mathbb{C}^m$ such that,

$$\forall z \in \mathbb{C}^m, \quad \text{proj}_{\mathcal{E}}(z) \in \arg\min_{h \in \mathcal{E}} \|h - z\|_2.$$

We call $\text{proj}_{\mathcal{E}}$ a *projection onto* \mathcal{E} .

We define $\text{proj}_{\text{Range}(\mathcal{A})}$ the standard orthogonal projection onto $\text{Range}(\mathcal{A})$.

The alternating projections algorithm, introduced in [Gerchberg and Saxton, 1972], addresses Problem (Set intersection) as follows: it starts at an arbitrary point $z_0 \in \mathbb{C}^m$, and iteratively defines, for all $t \in \mathbb{N}$,

$$z_{t+1} = \text{proj}_{\text{Range}(\mathcal{A})} \circ \text{proj}_{\mathcal{E}}(z_t).$$

3. Show that the sequence of iterates $(z_t)_{t \in \mathbb{N}}$ is bounded and satisfies

$$\|\text{proj}_{\mathcal{E}}(z_{t+1}) - z_{t+1}\|_2 \leq \|\text{proj}_{\mathcal{E}}(z_t) - z_t\|_2, \quad \forall t \in \mathbb{N}^*.$$

This property does not guarantee that $(z_t)_{t \in \mathbb{N}}$ converges towards a solution of (Set intersection) and, indeed, convergence does not always occur. However, it occurs sufficiently often so that alternating projections are one of the most standard phase retrieval algorithms.

Exercise 12: rank 1 approximation (local convergence)

This exercise and the next one are inspired by [Chi, Lu, and Chen, 2019]. Let $M \in \mathbb{R}^{d \times d}$ be a positive semidefinite matrix. We consider the problem of finding the rank 1 matrix which best approximates M in Frobenius norm. As any semidefinite matrix with rank at most 1 can be written as xx^T for some vector $x \in \mathbb{R}^d$, this amounts to finding a minimizer of

$$f : x \in \mathbb{R}^d \rightarrow \frac{1}{4} \|xx^T - M\|_F^2.$$

(The constant $\frac{1}{4}$ is only here to make formulas slightly nicer.)

1. Let $\lambda_1, \lambda_2, \dots, \lambda_d \geq 0$ be the eigenvalues of M , sorted in nonincreasing order. We assume that $1 = \lambda_1 > \lambda_2$.

Let (u_1, \dots, u_d) be an orthonormal basis of eigenvectors.

- a) Show that, for any x , $f(x) = \frac{1}{4} (\|x\|_2^4 - 2 \langle x, Mx \rangle + \|M\|_F^2)$.
- b) Show that f has at least one minimizer.
- c) Show that, for all $x \in \mathbb{R}^d$,

$$\nabla f(x) = \|x\|_2^2 x - Mx.$$

- d) Show that the minimizers of f are u_1 and $-u_1$.

We imagine that we run gradient descent on f , with stepsize $\tau \leq \frac{1}{2}$, starting at a point $x_0 \in \mathbb{R}^d$ such that

$$\|x_0 - u_1\|_2 < \frac{1 - \lambda_2}{7}.$$

It yields a sequence of iterates $(x_t)_{t \in \mathbb{N}}$. We are going to show that it converges to u_1 exponentially fast, more precisely that, for all $t \in \mathbb{N}$,

$$\|x_t - u_1\|_2 \leq \left(1 - \frac{(1 - \lambda_2)\tau}{2}\right)^t \|x_0 - u_1\|_2. \quad (4)$$

For all t , we define $\alpha_t \in \mathbb{R}, v_t \in \mathbb{R}^d$ such that

$$x_t = \alpha_t u_1 + v_t \quad \text{and} \quad v_t \in \text{Vect}\{u_2, \dots, u_d\}.$$

2. For any t , express $\|x_t - u_1\|_2$ as a function of $|\alpha_t - 1|$ and $\|v_t\|_2$.

3. a) Show that, for any t ,

$$\begin{aligned}\alpha_{t+1} &= (1 + \tau)\alpha_t - \tau\alpha_t^3 - \tau\alpha_t\|v_t\|_2^2; \\ v_{t+1} &= (1 - \tau(\alpha_t^2 + \|v_t\|_2^2))v_t + \tau Mv_t.\end{aligned}$$

b) Show that the first of these equalities is equivalent to

$$\alpha_{t+1} - 1 = (1 - \tau\alpha_t(\alpha_t + 1))(\alpha_t - 1) - \tau\alpha_t\|v_t\|_2^2.$$

From now on, we assume that Inequality (4) is true up to some step t .

4. Show that $1 - (\frac{1-\lambda_2}{7}) \leq \alpha_t \leq 1 + (\frac{1-\lambda_2}{7})$ and $\|v_t\|_2 \leq \frac{1-\lambda_2}{7}$.
5. Using Question 3.b), show that

$$|\alpha_{t+1} - 1| \leq \left(1 - \frac{5}{7}(1 - \lambda_2)\tau\right)|\alpha_t - 1| + \frac{8}{49}(1 - \lambda_2)\tau\|v_t\|_2.$$

6. a) Using Question 3.a), show that

$$\|v_{t+1}\|_2 \leq (1 - \tau(\alpha_t^2 + \|v_t\|_2^2 - \lambda_2))\|v_t\|_2.$$

[Hint: decompose v_t onto the orthogonal basis (u_1, \dots, u_d) .]

- b) Show that $0 \leq 1 - \tau(\alpha_t^2 + \|v_t\|_2^2 - \lambda_2) \leq 1 - \frac{5}{7}(1 - \lambda_2)\tau$.
7. a) Combine Questions 5 and 6 and show that

$$\sqrt{|\alpha_{t+1} - 1|^2 + \|v_{t+1}\|_2^2} \leq \left(1 - \frac{(1 - \lambda_2)\tau}{2}\right)\sqrt{|\alpha_t - 1|^2 + \|v_t\|_2^2}.$$

b) Conclude.

Exercise 13: rank 1 approximation (global convergence)

We keep the notation of the previous exercise. In particular, we still consider the function

$$f : x \in \mathbb{R}^d \rightarrow \frac{1}{4}\|xx^T - M\|_F^2,$$

and still assume that $1 = \lambda_1 > \lambda_2$.

1. Show that, for any $x \in \mathbb{R}^d$,

$$\text{Hess}f(x) = \|x\|_2^2 I_d + 2xx^T - M.$$

2. a) Compute the first-order critical points of f .
b) Compute the second-order critical points of f .
3. Show that, for almost any x_0 , if we choose a small enough stepsize, the sequence of gradient descent iterates converges to a minimizer of f .

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 $v \in \mathbb{R}^d \rightarrow \|Av\|_2 \in \mathbb{R}$ 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 $v \in \mathbb{R}^d \rightarrow \|Av\|_2 \in \mathbb{R}$ 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 orthogonal 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 $v_{k'}$,

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

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

$$\|\cos(\theta)v_k + \sin(\theta)v_{k'}\|_2 = \sqrt{\cos^2(\theta)\|v_k\|_2^2 + \sin^2(\theta)\|v_{k'}\|_2^2} = 1. \quad (5)$$

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

$$\cos(\theta)v_k + \sin(\theta)v_{k'} \in \text{Vect}\{v_1, \dots, v_{k-1}\}^\perp. \quad (6)$$

Equations (5) and (6), together with the definition of v_k , imply:

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

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

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

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

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

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

- 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

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

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 $k = D, \dots, d$, 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,

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

which is equivalent to

$$AV^T = UD,$$

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

$$A = UDV.$$

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

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

From the definition of the SVD,

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

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 $\text{Range}(A)$.

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

$$\begin{aligned} UDVx_* &= Ax_* = y \quad \text{and} \quad UDVx_\epsilon = Ax_\epsilon = y + \epsilon, \\ \Rightarrow D(Vx_*) &= U^T y \quad \text{and} \quad D(Vx_\epsilon) = U^T(y + \epsilon) = U^T y + U^T \epsilon. \end{aligned} \tag{7}$$

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 (7), for all $k \leq d$,

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

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

$$\begin{aligned}
\|V(x_* - x_\epsilon)\|_2 &= \left(\sum_{k=1}^d \left(x_{V,k} - x_{V,k}^{(\epsilon)} \right)^2 \right)^{1/2} \\
&= \left(\sum_{k=1}^d \frac{\epsilon_{U,k}^2}{\lambda_k^2} \right)^{1/2} \leq \left(\sum_{k=1}^d \frac{\epsilon_{U,k}^2}{\lambda_d^2} \right)^{1/2} \\
&= \frac{1}{\lambda_d} \left(\sum_{k=1}^m \epsilon_{U,k}^2 \right)^{1/2} = \frac{\|U^T \epsilon\|_2}{\lambda_d}.
\end{aligned}$$

Therefore,

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

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 :

$$y = Ue_1, \quad \epsilon = Ue_d,$$

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 $z \in \text{Ker}(A)$ 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.

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

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 $\|x_{kS} - x_*\|_2 \xrightarrow{k \rightarrow +\infty} \ell(x_*)$ and $\|x_{kS+1} - x_*\|_2 \xrightarrow{k \rightarrow +\infty} \ell(x_*)$, 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 $x_k \xrightarrow{k \rightarrow +\infty} x_\infty$.

Answer of Exercise 5

1. a) We set $S = \{s \leq m, \langle x, v_s \rangle = \langle x', v_s \rangle\}$. From the complement property, either $\text{Vect}\{v_s\}_{s \in \S} = \mathbb{R}^d$ or $\text{Vect}\{v_s\}_{s \notin S} = \mathbb{R}^d$.
First case: $\text{Vect}\{v_s\}_{s \in \S} = \mathbb{R}^d$. For any $s \in S$,

$$\langle x - x', v_s \rangle = \langle x, v_s \rangle - \langle x', v_s \rangle = 0.$$

Consequently, $x - x' \in \text{Vect}\{v_s\}_{s \in \S}^\perp = \{0\}$, meaning that $x = x'$.

Second case: $\text{Vect}\{v_s\}_{s \notin S} = \mathbb{R}^d$. For any s , $|\langle x, v_s \rangle| = |\langle x', v_s \rangle|$, hence $\langle x, v_s \rangle = \pm \langle x', v_s \rangle$. For any $s \notin S$, $\langle x, v_s \rangle \neq \langle x', v_s \rangle$, so that $\langle x, v_s \rangle = -\langle x', v_s \rangle$, and

$$\langle x + x', v_s \rangle = 0.$$

Consequently, $x + x' \in \text{Vect}\{v_s\}_{s \notin S}^\perp = \{0\}$, meaning that $x = -x'$.

- b) Let $S \subset \{1, \dots, m\}$ be such that

$$\text{Vect}\{v_s\}_{s \in \S} \neq \mathbb{R}^d \text{ and } \text{Vect}\{v_s\}_{s \notin \S} \neq \mathbb{R}^d.$$

Then

$$\text{Vect}\{v_s\}_{s \in \S}^\perp \neq \{0\} \text{ and } \text{Vect}\{v_s\}_{s \notin \S}^\perp \neq \{0\}.$$

Let $z_1 \in \text{Vect}\{v_s\}_{s \in \S}^\perp$ and $z_2 \in \text{Vect}\{v_s\}_{s \notin \S}^\perp$ be two non-zero vectors. For each $s \leq m$,

- if $s \in S$, then $\langle z_1, v_s \rangle = 0$,
 - if $s \notin S$, then $\langle z_2, v_s \rangle = 0$.
- c) For all $s \leq m$, $|\langle x, v_s \rangle| = |\langle x', v_s \rangle|$. Indeed, either $\langle z_1, v_s \rangle = 0$, in which case

$$\langle x, v_s \rangle = \langle z_2, v_s \rangle = -\langle x', v_s \rangle,$$

or $\langle z_2, v_s \rangle = 0$, in which case

$$\langle x, v_s \rangle = \langle z_1, v_s \rangle = \langle x', v_s \rangle.$$

Therefore, x and x' are two solutions of Problem (Real-PR) for $y = (|\langle x, v_s \rangle|)_{s \in S}$. But $x \neq x'$, since $z_2 \neq 0$, and $x \neq -x'$, since $z_1 \neq 0$.

2. a) Let us assume $m \leq 2d - 2$. Let $S \subset \{1, \dots, m\}$ have cardinality $\min(d - 1, m)$. It must hold $\text{Vect}\{v_s\}_{s \in S} \neq \mathbb{R}^d$, since $\text{Vect}\{v_s\}_{s \in S}$ has dimension at most $\text{Card}(S) \leq d - 1$.

On the other hand, $\{1, \dots, m\} \setminus S$ has cardinality $m - \text{Card}(S) = \max(0, m - d + 1) \leq d - 1$. It must therefore also hold $\text{Vect}\{v_s\}_{s \notin S} \neq \mathbb{R}^d$.

This shows that the complement property does not hold. From Question 1., Problem (Real-PR) has a non-unique solution for some values of y_1, \dots, y_m .

- b) For each n , let λ_n denote the Lebesgue measure over \mathbb{R}^n . For any $A \subset \{1, \dots, m\}$ with cardinality at most $d - 1$ and each $a \in \{1, \dots, m\} \setminus A$, we define

$$\mathcal{E}_{A,a} = \{(v_1, \dots, v_m) \in (\mathbb{R}^d)^m, v_a \in \text{Vect}\{v_s\}_{s \in A}\}.$$

Observe that, for each A, a ,

Answer of Exercise 6

1. The vector x_* is feasible for the problem (Basis Pursuit): $Ax_* = y$. Therefore, its ℓ^1 -norm is at least as large as the optimal value of the problem:

$$\|x_*\|_1 \geq \|x_{BP}\|_1 = \|x_* + h\|_1.$$

As a consequence,

$$\sum_{i \in T_*} |x_{*i}| = \|x_*\|_1$$

$$\begin{aligned}
&\geq \|x_* + h\|_1 \\
&= \sum_i |(x_* + h)_i| \\
&= \sum_{i \in T_*} |x_{*i} + h_i| + \sum_{i \notin T_*} |h_i| \\
&\geq \sum_{i \in T_*} (|x_{*i}| - |h_i|) + \sum_{i \notin T_*} |h_i| \\
&= \left(\sum_{i \in T_*} |x_{*i}| \right) - \|h_{T_*}\|_1 + \|h_{T_*^c}\|_1.
\end{aligned}$$

This implies $\|h_{T_*}\|_1 \geq \|h_{T_*^c}\|_1$.

2. a) For any $s \in T_l, s' \in T_{l-1}$, because the coordinates of h are non-increasing outside T_* ,

$$|h_{s'}| \geq |h_s|.$$

This implies that, for any $s \in T_l$,

$$\|h_{T_{l-1}}\|_1 = \sum_{s' \in T_{l-1}} |h_{s'}| \geq (\text{Card}(T_{l-1}))|h_s| = 3k|h_s|.$$

From this, we deduce that

$$\begin{aligned}
\|h_{T_l}\|_2^2 &= \sum_{s \in T_l} |h_s|^2 \\
&\leq \sum_{s \in T_l} \frac{\|h_{T_{l-1}}\|_1^2}{(3k)^2} \\
&= \frac{\|h_{T_{l-1}}\|_1^2}{(3k)^2} (\text{Card}(T_l)) \\
&\leq \frac{\|h_{T_{l-1}}\|_1^2}{3k}.
\end{aligned}$$

b)

$$\begin{aligned}
\sum_{l=2}^L \|h_{T_l}\|_2 &\leq \frac{1}{\sqrt{3k}} \sum_{l=1}^{L-1} \|h_{T_l}\|_1 \quad \text{from the previous question} \\
&\leq \frac{1}{\sqrt{3k}} \sum_{l=1}^L \|h_{T_l}\|_1
\end{aligned}$$

$$\begin{aligned}
&= \frac{\|h_{T_*^c}\|_1}{\sqrt{3k}} \\
&\leq \frac{\|h_{T_*}\|_1}{\sqrt{3k}} \quad \text{from the first question.}
\end{aligned}$$

c) By Cauchy-Schwarz,

$$\|h_{T_*}\|_1 \leq \sqrt{\text{Card}(T_*)} \|h_{T_*}\|_2 \leq \sqrt{k} \|h_{T_*}\|_2.$$

Combined with the previous question, it yields

$$\sum_{l=2}^L \|h_{T_l}\|_2 \leq \frac{\|h_{T_*}\|_2}{\sqrt{3}}.$$

3. a) As x_{BP} is a feasible point of Problem (**Basis Pursuit**), we have $Ax_{BP} = y = Ax_* = A(x_{BP} - h) = Ax_{BP} - Ah$. Consequently, $Ah = 0$.
b) As $h = h_{T_* \cup T_1} + h_{T_2} + \dots + h_{T_L}$, we have

$$\|Ah\|_2 = \|Ah_{T_* \cup T_1} + Ah_{T_2} + \dots + Ah_{T_L}\|_2 \geq \|Ah_{T_* \cup T_1}\|_2 - \sum_{l=2}^L \|Ah_{T_l}\|_2.$$

The vector $h_{T_* \cup T_1}$ has at most $\text{Card}(T_*) + \text{Card}(T_1) \leq k + 3k = 4k$ non-zero coordinates. From the definition of the restricted isometry constant,

$$\|Ah_{T_* \cup T_1}\|_2 \geq (1 - \delta_{4k}) \|h_{T_* \cup T_1}\|_2.$$

Similarly, for any $l \in \{2, \dots, L\}$,

$$\|Ah_{T_l}\|_2 \leq (1 + \delta_{4k}) \|h_{T_l}\|_2.$$

This gives the desired inequality:

$$\|Ah\|_2 \geq (1 - \delta_{4k}) \|h_{T_* \cup T_1}\|_2 - (1 + \delta_{4k}) \sum_{l=2}^L \|h_{T_l}\|_2.$$

c) Together, the previous two subquestions imply

$$(1 - \delta_{4k}) \|h_{T_* \cup T_1}\|_2 \leq (1 + \delta_{4k}) \sum_{l=2}^L \|h_{T_l}\|_2$$

Using also Question 2.c,

$$\begin{aligned} (1 - \delta_{4k}) \|h_{T_*}\|_2 &\leq (1 - \delta_{4k}) \|h_{T_* \cup T_1}\|_2 \\ &\leq (1 + \delta_{4k}) \frac{\|h_{T_*}\|_2}{\sqrt{3}}. \end{aligned}$$

Since $\delta_{4k} < 1/4$, this implies

$$\frac{3}{4} \|h_{T_*}\|_2 \leq \frac{5}{4} \frac{\|h_{T_*}\|_2}{\sqrt{3}} \Rightarrow \frac{3\sqrt{3}}{5} \|h_{T_*}\|_2 \leq \|h_{T_*}\|_2.$$

Since $\frac{3\sqrt{3}}{5} > 1$, this implies $\|h_{T_*}\|_2 = 0$: the coordinates of h with indices in T_* are zero. From the first question, the coordinates of h with indices in T_*^c are therefore also zero, so $h = 0$ and

$$x_{BP} = x_*.$$

Answer of Exercise 7

1. Since all non-zero columns of X_* are linearly independent, the rank of X_* is the number of non-zero columns, that is the number of non-zero λ_s . Since $\text{rank}(X_*) \leq r$, at most the r largest singular values are non-zero.
2. The matrix $X_* + H_c$ is diagonal. The coefficients on its diagonal (which are its singular values, since they are nonnegative) are $\lambda_1, \dots, \lambda_r, \mu_1, \dots, \mu_{d_2-r}$. Therefore,

$$\|X_* + H_c\|_* = \sum_{s=1}^r \lambda_s + \sum_{s=1}^{d_2-r} \mu_s = \|X_*\|_* + \|H_c\|_*.$$

In addition, since X_{NM} is a solution of Problem (Nuclear-min) and X_* is a feasible point of this problem,

$$\begin{aligned} \|X_*\|_* &\geq \|X_{NM}\|_* \\ &= \|X_* + H\|_* \\ &= \|X_* + H_c + H_0\|_* \\ &\geq \|X_* + H_c\|_* - \|H_0\|_* \\ &= \|X_*\|_* + \|H_c\|_* - \|H_0\|_*. \end{aligned}$$

Therefore, $\|H_0\|_* \geq \|H_c\|_*$.

3. a) For any $l \in \{2, \dots, L\}$,

$$\begin{aligned}
\|H_{c,l}\|_F &= \left(\sum_{s=3(l-1)r}^{3lr} \mu_s^2 \right)^{1/2} \\
&\leq \left(\sum_{s=3(l-1)r+1}^{3lr} \left(\frac{1}{3r} \sum_{t=3(l-2)r+1}^{3(l-1)r} \mu_s \right)^2 \right)^{1/2} \\
&= \left(\frac{1}{3r} \left(\sum_{t=3(l-2)r+1}^{3(l-1)r} \mu_s \right)^2 \right)^{1/2} \\
&= \frac{\|H_{c,l-1}\|_*}{\sqrt{3r}}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\sum_{l=2}^L \|H_{c,l}\|_F &\leq \sum_{l=2}^L \frac{\|H_{c,l-1}\|_*}{\sqrt{3r}} \\
&\leq \sum_{l=1}^L \frac{\|H_{c,l}\|_*}{\sqrt{3r}} \\
&= \frac{\|H_c\|_*}{\sqrt{3r}} \\
&\leq \frac{\|H_0\|_*}{\sqrt{3r}}.
\end{aligned}$$

b) The rank of H_0 is the dimension of the vector space generated by its columns. The first r columns, which form the matrix $\begin{pmatrix} H_{11} \\ H_{21} \end{pmatrix}$, generate a space with dimension at most r . The last $d_2 - r$ columns, which form the matrix $\begin{pmatrix} H_{12} \\ 0 \end{pmatrix}$, generate a vector space which is included in $\mathbb{R}^r \times \{0\}^{d_1-r}$, and therefore has dimension at most r . Therefore, the two sets of columns generate a set with dimension at most $r + r = 2r$, hence

$$\text{rank}(H_0) \leq 2r.$$

Let us denote $\alpha_1, \dots, \alpha_{d_2}$ the singular values of H_0 , ordered by decreasing order. Then, for the same reason as in Question 1., $\alpha_{2r+1} =$

$\dots = \alpha_{d_2} = 0$, so that

$$\begin{aligned} \|H_0\|_* &= \sum_{s=1}^{2r} \alpha_s \\ &\leq \sqrt{2r} \left(\sum_{s=1}^{2r} \alpha_s^2 \right)^{1/2} \quad \text{by Cauchy-Schwarz} \\ &= \sqrt{2r} \|H_0\|_F. \end{aligned}$$

[Remark: since multiplying a matrix with an orthogonal matrix does not change its Frobenius norm, the Frobenius norm of any matrix is equal to the Frobenius norm of the diagonal part of its singular value decomposition. It is therefore the ℓ^2 -norm of its singular values.]

$$\text{c) } \sum_{l=2}^L \|H_{c,l}\|_F \leq \frac{\|H_0\|_*}{\sqrt{3r}} \leq \frac{\sqrt{2r}\|H_0\|_F}{\sqrt{3r}} = \sqrt{\frac{2}{3}} \|H_0\|_*.$$

4. a)

$$\begin{aligned} \|\mathcal{L}(H)\|_2 &= \left\| \mathcal{L} \left(H_0 + H_{c,1} + \sum_{l=2}^L H_{c,l} \right) \right\|_2 \\ &\geq \|\mathcal{L}(H_0 + H_{c,1})\|_2 - \sum_{l=2}^L \|\mathcal{L}(H_{c,l})\|_2 \\ &\geq (1 - \delta_{5r}) \|H_0 + H_{c,1}\|_F - (1 + \delta_{5r}) \sum_{l=2}^L \|H_{c,l}\|_F. \end{aligned}$$

We have used the definition of the restricted isometry constant, the fact that $\text{rank}(H_{c,l}) \leq 3r$ for any l , and that $\text{rank}(H_0 + H_{c,1}) \leq \text{rank}(H_0) + \text{rank}(H_{c,1}) \leq 2r + 3r = 5r$.

b) From the previous two questions,

$$\begin{aligned} \|\mathcal{L}(H)\|_2 &\geq (1 - \delta_{5r}) \|H_0 + H_{c,1}\|_F - (1 + \delta_{5r}) \sqrt{\frac{2}{3}} \|H_0\|_F \\ &\geq (1 - \delta_{5r}) \|H_0\|_F - (1 + \delta_{5r}) \sqrt{\frac{2}{3}} \|H_0\|_F \\ &\geq \left(\frac{9}{10} - \frac{11}{10} \sqrt{\frac{2}{3}} \right) \|H_0\|_F. \end{aligned}$$

For the second inequality, we have use the facts that the non-zero coefficients of H_0 and $H_{c,1}$ are at different positions, hence $\|H_0 + H_{c,1}\|_F = \sqrt{\|H_0\|_F^2 + \|H_{c,1}\|_F^2} \geq \|H_0\|_F$.

As $\mathcal{L}(H) = 0$ (X_{NM} and X_0 are feasible for Problem (Nuclear-min)) and $\frac{9}{10} - \frac{11}{10}\sqrt{\frac{2}{3}} > 0$, this implies $\|H_0\|_F = 0$.

As a consequence, $H_0 = 0$. From Question 2., $H_c = 0$. This means that $H = 0$ and $X_{NM} = X_*$.

Answer of Exercise 9

1.

$$\begin{aligned} & \left| \sum_{k=-N}^N z_k e^{2\pi i k t} \right| \leq 1, \quad \forall t \in \mathbb{R}, \\ \iff & \left| \sum_{k=-N}^N z_k e^{2\pi i k t} \right|^2 \leq 1, \quad \forall t \in \mathbb{R}, \\ \iff & 1 - \left| \sum_{k=-N}^N z_k e^{2\pi i k t} \right|^2 \in \mathbb{R}^+ \quad \forall t \in \mathbb{R}. \end{aligned}$$

From Fejér-Riesz' theorem, this property holds if and only if there exists trigonometric polynomials P_1, \dots, P_n with degree at most N such that, for all $t \in \mathbb{R}$,

$$1 - \left| \sum_{k=-N}^N z_k e^{2\pi i k t} \right|^2 = \sum_{l=1}^n |P_l(e^{2\pi i t})|^2,$$

which is equivalent to $\left| \sum_{k=-N}^N z_k e^{2\pi i k t} \right|^2 + \sum_{l=1}^n |P_l(e^{2\pi i t})|^2 = 1$.

2. Let a_{-2N}, \dots, a_{2N} denote the coefficients of the polynomial in Equality (3):

$$\begin{aligned} \sum_{d=-2N}^{2N} a_d e^{2\pi i d t} &= \left| \sum_{k=-N}^N z_k e^{2\pi i k t} \right|^2 + \sum_{l=1}^n |P_l(e^{2\pi i t})|^2 \\ &= \left(\sum_{k=-N}^N z_k e^{2\pi i k t} \right) \left(\sum_{k=-N}^N \overline{z_k} e^{-2\pi i k t} \right) \end{aligned}$$

$$\begin{aligned}
& + \sum_{l=1}^n \left(\sum_{k=-N}^N p_k^{(l)} e^{2\pi i k t} \right) \left(\sum_{k=-N}^N \overline{p_k^{(l)}} e^{-2\pi i k t} \right) \\
& = \sum_{k=-N}^N \sum_{k'=-N}^N \left(z_k \overline{z_{k'}} + \sum_{l=1}^n p_k^{(l)} \overline{p_{k'}^{(l)}} \right) e^{2\pi i (k-k')t} \\
& = \sum_{d=-2N}^{2N} \sum_{\substack{-N \leq k, k' \leq N \\ k-k'=d}} \left(z_k \overline{z_{k'}} + \sum_{l=1}^n p_k^{(l)} \overline{p_{k'}^{(l)}} \right) e^{2\pi i d t} \\
& = \sum_{d=-2N}^{2N} \sum_{\substack{-N \leq k \leq N \\ \text{s.t. } -N \leq k-d \leq N}} \left(z_k \overline{z_{k-d}} + \sum_{l=1}^n p_k^{(l)} \overline{p_{k-d}^{(l)}} \right) e^{2\pi i d t} \\
& = \sum_{d=-2N}^{2N} \sum_{k=-N+\max(0,d)}^{N-\max(0,-d)} \left(z_k \overline{z_{k-d}} + \sum_{l=1}^n p_k^{(l)} \overline{p_{k-d}^{(l)}} \right) e^{2\pi i d t} \\
& = \sum_{d=-2N}^{2N} \left(\sum_{k=-N+\max(0,d)}^{N-\max(0,-d)} A_{k+N+1, k+N+1-d} \right) e^{2\pi i d t} \\
& = \sum_{d=-2N}^{2N} \left(\sum_{k=1+\max(0,d)}^{2N+1-\max(0,-d)} A_{k, k-d} \right) e^{2\pi i d t}.
\end{aligned}$$

As a consequence, $a_d = \sum_{k=1+\max(0,d)}^{2N+1-\max(0,-d)} A_{k, k-d}$ for all $d = -N, \dots, N$, and Equality (3) holds if and only if, for any $d = -N, \dots, N$,

$$\begin{aligned}
a_d &= \sum_{k=1+\max(0,d)}^{2N+1-\max(0,-d)} A_{k, k-d} = 1 \text{ if } d = 0, \\
&= 0 \text{ otherwise.}
\end{aligned}$$

3. If $A = zz^* + \sum_{l=1}^n p^{(l)} p^{(l)*}$ for some vectors $p^{(1)}, \dots, p^{(n)} \in \mathbb{C}^{2N+1}$, then $A - zz^* = \sum_{l=1}^n p^{(l)} p^{(l)*}$, which is semidefinite positive:

$$\forall x \in \mathbb{C}^{2N+1}, \quad x^*(A - zz^*)x = \sum_{l=1}^n |\langle p^{(l)}, x \rangle|^2 \geq 0.$$

Conversely, let us assume that $A - zz^* \succeq 0$. Let $B \in \mathbb{C}^{(2N+1) \times (2N+1)}$ be

a square root of $A - zz^*$ (that is, a matrix such that $BB^* = A - zz^*$).¹ Let $p^{(1)}, \dots, p^{(2N+1)}$ be the column vectors of B . Then

$$A - zz^* = \sum_{l=1}^{2N+1} p^{(l)} p^{(l)*},$$

which implies $A = zz^* + \sum_{l=1}^n p^{(l)} p^{(l)*}$ for $n = 2N + 1$.

4. Let us denote

$$G = \left(\begin{array}{c|c} A & z \\ \hline z^* & 1 \end{array} \right) \in \mathbb{C}^{(2N+2) \times (2N+2)}$$

and show, as required, that $A - zz^* \succeq 0$ if and only if $G \succeq 0$.

$$\begin{aligned} (G \succeq 0) &\iff (\forall h \in \mathbb{C}^{2N+2}, h^* G h \geq 0) \\ &\iff \left(\forall \tilde{h} \in \mathbb{C}^{2N+1}, u \in \mathbb{C}, \begin{pmatrix} \tilde{h} \\ u \end{pmatrix}^* G \begin{pmatrix} \tilde{h} \\ u \end{pmatrix} \geq 0 \right) \\ &\iff \left(\forall \tilde{h} \in \mathbb{C}^{2N+1}, u \in \mathbb{C}, \tilde{h}^* A \tilde{h} + 2\operatorname{Re}(u \langle \tilde{h}, z \rangle) + |u|^2 \geq 0 \right) \\ &\iff \left(\forall \tilde{h} \in \mathbb{C}^{2N+1}, t \in \mathbb{R}, \phi \in \mathbb{R}, \right. \\ &\quad \left. \tilde{h}^* A \tilde{h} + 2\operatorname{Re}(te^{i\phi} \langle \tilde{h}, z \rangle) + |te^{i\phi}|^2 \geq 0 \right) \\ &\iff \left(\forall \tilde{h} \in \mathbb{C}^{2N+1}, t \in \mathbb{R}, \phi \in \mathbb{R}, \right. \\ &\quad \left. \tilde{h}^* A \tilde{h} + 2t\operatorname{Re}(e^{i\phi} \langle \tilde{h}, z \rangle) + t^2 \geq 0 \right) \\ &\stackrel{(a)}{\iff} \left(\forall \tilde{h} \in \mathbb{C}^{2N+1}, \phi \in \mathbb{R}, \tilde{h}^* A \tilde{h} - \left(\operatorname{Re}(e^{i\phi} \langle \tilde{h}, z \rangle) \right)^2 \geq 0 \right) \\ &\stackrel{(b)}{\iff} \left(\forall \tilde{h} \in \mathbb{C}^{2N+1}, \tilde{h}^* A \tilde{h} - |\langle \tilde{h}, z \rangle|^2 \geq 0 \right) \\ &\iff \left(\forall \tilde{h} \in \mathbb{C}^{2N+1}, \tilde{h}^* (A - zz^*) \tilde{h} \geq 0 \right) \\ &\iff A - zz^* \succeq 0. \end{aligned}$$

Equivalence (a) is true because, for any \tilde{h} and ϕ , it holds that the polynomial $t \rightarrow \tilde{h}^* A \tilde{h} + 2t\operatorname{Re}(e^{i\phi} \langle \tilde{h}, z \rangle) + t^2$ is nonnegative over \mathbb{R} if and

¹All semidefinite positive matrices have square roots; it is most easily proved by writing the semidefinite matrix in an eigenvector basis.

only if its discriminant is nonpositive:

$$4 \left(\operatorname{Re} \left(e^{i\phi} \langle \tilde{h}, z \rangle \right) \right)^2 - 4 \tilde{h}^* A \tilde{h} \leq 0,$$

which is exactly $\tilde{h}^* A \tilde{h} - \left(\operatorname{Re} \left(e^{i\phi} \langle \tilde{h}, z \rangle \right) \right)^2 \geq 0$.

Equivalence (b) is true because, for any \tilde{h} , we have $\tilde{h}^* A \tilde{h} - \left(\operatorname{Re} \left(e^{i\phi} \langle \tilde{h}, z \rangle \right) \right)^2 \geq 0$ for all $\phi \in \mathbb{R}$ if and only if the minimum over ϕ of this quantity is non-negative, and the minimum is precisely

$$\tilde{h}^* A \tilde{h} - |\langle \tilde{h}, z \rangle|^2.$$

5. Since both problems have the same objective function, it suffices to show that z is feasible for **(Dual TV)** if and only if z is feasible for the other problem, that is there exists $A \in \mathbb{C}^{(2N+1) \times (2N+1)}$ such that

$$\begin{aligned} \sum_{k=1+\max(0,d)}^{2N+1-\max(0,-d)} A_{k,k-d} &= 0 \text{ for all } d \in \{-2N, \dots, 2N\} \setminus \{0\}, \\ \sum_{k=1}^{2N+1} A_{k,k} &= 1, \\ \text{and } \left(\begin{array}{c|c} A & z \\ \hline z^* & 1 \end{array} \right) &\succeq 0. \end{aligned}$$

Let us first assume that z is feasible for **(Dual TV)**: $\left| \sum_{k=-N}^N z_k e^{2\pi i k t} \right| \leq 1$ for all $t \in \mathbb{R}$. From Question 1, there exists trigonometric polynomials with degree at most N , P_1, \dots, P_n , such that

$$\left| \sum_{k=-N}^N z_k e^{2\pi i k t} \right|^2 + \sum_{l=1}^n |P_l(e^{2\pi i t})|^2 = 1.$$

We denote $p^{(1)}, \dots, p^{(n)}$ the vectors of their coefficients and set $A = z z^* + \sum_{l=1}^n p^{(l)} p^{(l)*}$. From Question 2, we have

$$\sum_{k=1+\max(0,d)}^{2N+1-\max(0,-d)} A_{k,k-d} = 0 \text{ for all } d \in \{-2N, \dots, 2N\} \setminus \{0\},$$

$$\sum_{k=1}^{2N+1} A_{k,k} = 1.$$

From Question 3, $A - zz^* \succeq 0$, hence from Question 4,

$$\left(\begin{array}{c|c} A & z \\ \hline z^* & 1 \end{array} \right) \succeq 0.$$

The existence of A is proved.

Conversely, let us assume the existence of A , and show that $\left| \sum_{k=-N}^N z_k e^{2\pi i k t} \right| \leq 1$ for all $t \in \mathbb{R}$. From Question 4, $A - zz^* \succeq 0$. Therefore, from Question 3, there exist $p^{(1)}, \dots, p^{(n)}$ such that

$$A = zz^* + \sum_{l=1}^n p^{(l)} p^{(l)*}.$$

Let us denote P_1, \dots, P_n the corresponding trigonometric polynomials. From Question 2, they satisfy Equality (3)

$$\left| \sum_{k=-N}^N z_k e^{2\pi i k t} \right|^2 + \sum_{l=1}^n |P_l(e^{2\pi i t})|^2 = 1, \quad \forall t \in \mathbb{R}.$$

From Question 1, this means that $\left| \sum_{k=-N}^N z_k e^{2\pi i k t} \right| \leq 1$ for all $t \in \mathbb{R}$.

Answer of Exercise 10

1. It is a non-convex problem, because \mathcal{E}_k is not-convex. Indeed, \mathcal{E}_k contains $\mathbb{R}e_E$ for all $E \subset \{1, \dots, d\}$, therefore $\text{Conv}(\mathcal{E}_k) \supset \text{Vect}(\{e_E\}_{E \subset \{1, \dots, d\}}) = \mathbb{R}^d$. But $\mathcal{E}_k \neq \mathbb{R}^d$ (it is a finite union of k -dimensional vector subspaces of \mathbb{R}^d , and $k < d$), so \mathcal{E}_k is different from its convex hull.
2. First, let x be any point of this set, which we denote \mathcal{M} . We assume that $x \neq e_E$ for all E and show that x is not extremal. For all i , $x_i \in [-1; 1]$. Since $x \neq e_E$ for all E , there exists an index i such that $x_i \neq -1$ and $x_i \neq 1$ (i.e. $x_i \in]-1; 1[$). We define $y, z \in \mathbb{R}^d$ such that

$$\begin{aligned} y_j &= z_j = x_j, \forall j \neq i, \\ y_i &= 1, \end{aligned}$$

$$z_i = -1.$$

Then $x = (1 - t)y + tz$ for $t = \frac{1-x_i}{2} \in [0; 1]$. The vectors y, z belong to \mathcal{M} and are different from x . Therefore, x is not extremal.

Now, let us fix $E \subset \{1, \dots, d\}$. The vector e_E is in \mathcal{M} . Let us show that it is extremal. Let $y, z \in \mathcal{M}$ and $t \in [0; 1]$ be such that $e_E = (1 - t)y + tz$. We must show that either y or z is equal to e_E .

If $t = 0$, then $y = e_E$. If $t = 1$, then $z = e_E$. Let us assume $0 < t < 1$ and show that $y = z = e_E$. For all $i \leq d$, if $i \in E$, then $(e_E)_i = 1$. Since $y_i \leq 1$ and $z_i \leq 1$, it holds

$$1 = (e_E)_i = (1 - t)y_i + tz_i \leq (1 - t) + t = 1.$$

The inequality must be an equality, meaning that $y_i = z_i = 1 = (e_E)_i$. The same reasoning applies if $i \notin E$. It shows that y, z and e_E are equal.

3. The problem we want to solve is (assuming that a solution exists)

$$\begin{aligned} & \text{minimize } \|x\|_{reg} \\ & \text{over } x \in \mathbb{R}^d \text{ such that } Ax = y, \end{aligned}$$

where, for any x , $\|x\|_{reg} = \min\{s \in \{1, \dots, d\}, x \in \mathcal{E}_s\}$.

Following the intuition discussed in the lectures, since the vectors e_E are the extremal points of the ℓ^∞ ball, it makes sense to approximate $\|\cdot\|_{reg}$ with the infinity ball, yielding the convex problem

$$\begin{aligned} & \text{minimize } \|x\|_\infty \\ & \text{over } x \in \mathbb{R}^d \\ & \text{such that } Ax = y. \end{aligned} \tag{Relax-Reg}$$

4. Mimicking the definition of k -restricted isometry constant for sparse recovery, we could define the k -restricted isometry constant of A as the smallest positive number $\delta_k > 0$ (if it exists) such that, for all $x \in \mathcal{E}_k$,

$$(1 - \delta_k)\|x\|_2 \leq \|Ax\|_2 \leq (1 + \delta_k)\|x\|_2.$$

5. a) Problem (Relax-Reg) can be rewritten as a min-max problem:

$$\min_{x \in \mathbb{R}^d} \max_{z \in \mathbb{R}^m} \|x\|_\infty + \langle y - Ax, z \rangle$$

$$= \min_{x \in \mathbb{R}^d} \max_{z \in \mathbb{R}^m} \langle y, z \rangle + \|x\|_\infty - \langle x, A^T z \rangle$$

We compute the dual by exchanging the minimum and maximum:

$$\begin{aligned} & \max_{z \in \mathbb{R}^m} \min_{x \in \mathbb{R}^d} \langle y, z \rangle + \|x\|_\infty - \langle x, A^T z \rangle \\ &= \max_{z \in \mathbb{R}^m} \langle y, z \rangle + \min_{x \in \mathbb{R}^d} (\|x\|_\infty - \langle x, A^T z \rangle) \\ &= \max_{\substack{z \in \mathbb{R}^m \\ \|A^T z\|_1 \leq 1}} \langle y, z \rangle. \end{aligned}$$

Indeed, $\min_{x \in \mathbb{R}^d} \|x\|_\infty - \langle x, A^T z \rangle = -\infty$ if $\|A^T z\|_1 > 1$: denoting $h \in \mathbb{R}^d$ the vector such that $h_i = 1$ if $(A^T z)_i \geq 0$ and $h_i = -1$ if $(A^T z)_i < 0$, we have, for all $t \geq 0$,

$$\begin{aligned} \|th\|_\infty - \langle th, A^T z \rangle &= t - t\|A^T z\|_1 \\ &= -t(\|A^T z\|_1 - 1), \end{aligned}$$

which goes to $-\infty$ when t goes to ∞ .

On the other hand, if $\|A^T z\|_1 \leq 1$, then $\|x\|_\infty - \langle x, A^T z \rangle \geq \|x\|_\infty - \|x\|_\infty \|A^T z\|_1 \geq 0$ for all $x \in \mathbb{R}^d$. Since $x = 0$ yields the value 0, the minimum is zero.

The dual problem is

$$\begin{aligned} & \text{maximize } \langle y, z \rangle, \\ & \text{over } z \in \mathbb{R}^m \\ & \text{such that } \|A^T z\|_1 \leq 1. \end{aligned}$$

- b) First, we observe that $(x_o, z_o) \in \mathbb{R}^d \times \mathbb{R}^m$ is primal-dual optimal if and only if

$$\begin{aligned} \|x_o\|_\infty &= \max_{z \in \mathbb{R}^m} \langle y, z \rangle + \|x_o\|_\infty - \langle x_o, A^T z \rangle \\ &= \langle y, z_o \rangle + \|x_o\|_\infty - \langle x_o, A^T z_o \rangle \\ &= \min_{x \in \mathbb{R}^d} \langle y, z_o \rangle + \|x\|_\infty - \langle x, A^T z_o \rangle \\ &= \langle y, z_o \rangle. \end{aligned}$$

This is equivalent to

- x_o is primal feasible: $y = Ax_o$;
- $\|A^T z_o\|_1 = 1$;
- $(x_o)_i = \text{sgn}((A^T z_o)_i) \|x_o\|_\infty$ for all i such that $(A^T z_o)_i \neq 0$.

We observe that (x_o, c) is primal-dual optimal. Indeed, x_o is primal feasible and c is dual feasible. For any i , $(x_o)_i = \pm \|x_o\|_\infty$ (because $k = 1$, hence x_o is proportional to e_E for some set E). Therefore, from the definition of c , $(A^T c)_i \neq 0$ and $(A^T c)_i$ has the same sign as $(x_o)_i$, meaning that

$$(x_o)_i = \text{sgn}((A^T c)_i) \|x_o\|_\infty.$$

Therefore, x_o is a minimizer of (Relax-Reg).

Let us show that the minimizer is unique. Let x_* be any minimizer of (Relax-Reg). Then (x_*, c) is primal-dual optimal. Therefore, $y = Ax_*$ and

$$(x_*)_i = \text{sgn}((A^T c)_i) \|x_*\|_\infty \text{ for all } i \text{ such that } (A^T c)_i \neq 0.$$

This means that

$$\begin{aligned} (x_*)_i &= \|x_*\|_\infty, \forall i \leq d \text{ such that } (x_o)_i = \|x_o\|_\infty, \\ (x_*)_i &= -\|x_*\|_\infty, \forall i \leq d \text{ such that } (x_o)_i = -\|x_o\|_\infty. \end{aligned}$$

Consequently, x_* is proportional to x_o . Since $Ax_* = y = Ax_o$ and $y \neq 0$, the proportionality constant is equal to 1, meaning that $x_* = x_o$.

Answer of Exercise 11

2. Let z belong to \mathbb{C}^m . Let us determine $\text{argmin}_{h \in \mathcal{E}} \|h - z\|_2$. A vector h belongs to \mathcal{E} if and only if there exists $\phi_1, \dots, \phi_m \in \mathbb{R}$ such that

$$h_s = y_s e^{i\phi_s}, \quad \forall s = 1, \dots, m.$$

Therefore,

$$\begin{aligned} \min_{h \in \mathcal{E}} \|h - z\|_2 &= \min_{h \in \mathcal{E}} \left(\sum_{j=1}^m |h_j - z_j|^2 \right)^{1/2} \\ &= \min_{\phi_1, \dots, \phi_m \in \mathbb{R}} \left(\sum_{j=1}^m |y_j e^{i\phi_j} - z_j|^2 \right)^{1/2} \end{aligned}$$

$$= \min_{\phi_1, \dots, \phi_m \in \mathbb{R}} \left(\sum_{j=1}^m y_j^2 - 2\operatorname{Re}(y_j z_j e^{-i\phi_j}) + |z_j|^2 \right)^{1/2}.$$

This formula is minimized exactly when $-\operatorname{Re}(y_j z_j e^{-i\phi_j})$ is minimized for all $j = 1, \dots, m$. When $y_j = 0$ or $z_j = 0$, $-\operatorname{Re}(y_j z_j e^{-i\phi_j}) = 0$ for all real numbers ϕ_j , so all real numbers ϕ_j are minimizers. When $y_j, z_j \neq 0$, the minimum is attained when

$$\phi_j \equiv \operatorname{phase}(y_j z_j) \quad [2\pi],$$

which is equivalent to $\phi_j \equiv \operatorname{phase}(z_j) [2\pi]$, since y_j is a positive real number. To summarize, an element h of \mathcal{E} minimizes the distance to z if and only if, for all $j \leq m$,

$$\begin{aligned} h_j &= y_j e^{i\operatorname{phase}(z_j)} \text{ if } y_j, z_j \neq 0, \\ h_j &= 0 \text{ if } y_j = 0, \\ h_j &= y_j e^{i\phi_j} \text{ for some } \phi_j \in \mathbb{R} \text{ if } z_j = 0. \end{aligned}$$

Therefore, a possible expression for $\operatorname{proj}_{\mathcal{E}}$ is to define, for every $z \in \mathbb{C}^m$,

$$\operatorname{proj}_{\mathcal{E}}(z) = (y_j e^{i\operatorname{phase}(z_j)})_{j=1, \dots, m},$$

with the convention that $e^{i\operatorname{phase}(z_j)} = 1$ if $z_j = 0$.

Answer of Exercise 12

1. a) For any $x \in \mathbb{R}^d$,

$$\begin{aligned} f(x) &= \frac{1}{4} \|xx^T - M\|_F^2 \\ &= \frac{1}{4} \operatorname{Tr}((xx^T - M)(xx^T - M)^T) \\ &= \frac{1}{4} \operatorname{Tr}(xx^T xx^T - Mxx^T - xx^T M + MM^T) \\ &= \frac{1}{4} (\operatorname{Tr}(xx^T xx^T) - 2\operatorname{Tr}(Mxx^T) + \operatorname{Tr}(MM^T)) \\ &= \frac{1}{4} (\operatorname{Tr}(x^T xx^T x) - 2\operatorname{Tr}(x^T(Mx)) + \|M\|_F^2) \\ &= \frac{1}{4} (\|x\|_2^4 - 2\langle x, Mx \rangle + \|M\|_F^2). \end{aligned}$$

b) For any $x \in \mathbb{R}^d$, $\|Mx\|_2 \leq \lambda_1 \|x\|_2$, hence $|\langle x, Mx \rangle| \leq \lambda_1 \|x\|_2^2$, and

$$\begin{aligned} f(x) &\geq \frac{\|x\|_2^4}{4} - \frac{\lambda_1 \|x\|_2^2}{2} \\ &= \frac{\|x\|_2^2}{2} \left(\frac{\|x\|_2^2}{2} - \lambda_1 \right) \\ &\rightarrow +\infty \quad \text{when } \|x\|_2 \rightarrow +\infty. \end{aligned}$$

This shows that f is coercive. It is also continuous, hence has a minimizer.

c) For all $x, h \in \mathbb{R}^d$,

$$\begin{aligned} f(x+h) &= \frac{1}{4} (\|x+h\|_2^4 - 2\langle x+h, M(x+h) \rangle + \|M\|_F^2) \\ &= \frac{1}{4} (\|x\|_2^2 + 2\langle x, h \rangle + \|h\|_2^2)^2 \\ &\quad - 2(\langle x, Mx \rangle + \langle h, Mx \rangle + \langle x, Mh \rangle + \langle h, Mh \rangle) + \|M\|_F^2 \\ &= \frac{1}{4} (\|x\|_2^4 + 4\|x\|_2^2 \langle x, h \rangle - 2\langle x, Mx \rangle - 4\langle Mx, h \rangle + \|M\|_F^2 + o(\|h\|_2)) \\ &= f(x) + \langle \|x\|_2^2 x - Mx, h \rangle + o(\|h\|_2). \end{aligned}$$

Therefore, $\nabla f(x) = \|x\|_2^2 x - Mx$.

d) Let us first consider an arbitrary minimizer x_{min} . We must have

$$0 = \nabla f(x_{min}) = \|x_{min}\|_2^2 x_{min} - Mx_{min}.$$

As a consequence, $Mx_{min} = \|x_{min}\|_2^2 x_{min}$, which means that x_{min} is an eigenvector of M , with eigenvalue $\|x_{min}\|_2^2$. In particular, there exists $k = 1, \dots, d$ such that

- x_{min} is an eigenvector of M with eigenvalue λ_k ;
- $\|x_{min}\|_2^2 = \lambda_k$, that is, $\|x_{min}\|_2 = \sqrt{\lambda_k}$.

This shows that minimizers of f are necessarily of the form $x_{min} = \sqrt{\lambda_k} v$, for v a unitary eigenvector associated to the eigenvalue λ_k .

Now, we compute the minimizers. For $k \leq d$ and v as above,

$$\begin{aligned} f(\sqrt{\lambda_k} v) &= \frac{1}{4} \left(\|\sqrt{\lambda_k} v\|_2^4 - 2\langle \sqrt{\lambda_k} v, M\sqrt{\lambda_k} v \rangle + \|M\|_F^2 \right) \\ &= \frac{1}{4} (-\lambda_k^2 + \|M\|_F^2). \end{aligned}$$

This is minimal if and only if $\lambda_k = \lambda_1 (= 1)$ and v is an eigenvector associated to the eigenvalue λ_1 , that is to say $v = \pm u_1$. Therefore, the minimizers are u_1 and $-u_1$.

2. As $x_t - u_1 = (\alpha_t - 1)u_1 + v_t$ and $u_1 \perp v_t$, the norm is

$$\|x_t - u_1\|_2 = \sqrt{|\alpha_t - 1|^2 + \|v_t\|_2^2}.$$

3. a)

$$\begin{aligned} x_{t+1} &= x_t - \tau \nabla f(x_t) \\ &= \alpha_t u_1 + v_t - \tau(\|x_t\|_2^2 x_t - Mx_t) \\ &= \alpha_t u_1 + v_t - \tau(\|x_t\|_2^2 (\alpha_t u_1 + v_t) - \alpha_t M u_1 - M v_t) \\ &= \alpha_t(1 - \tau\|x_t\|_2^2)u_1 + (1 - \tau\|x_t\|_2^2)v_t + \tau\alpha_t u_1 + \tau M v_t \\ &= \alpha_t(1 - \tau(\alpha_t^2 + \|v_t\|_2^2) + \tau)u_1 + (1 - \tau(\alpha_t^2 + \|v_t\|_2^2))v_t + \tau M v_t. \end{aligned}$$

As v_t and Mv_t belong to $\text{Vect}\{u_2, \dots, u_d\}$,

$$\begin{aligned} \alpha_{t+1} &= \alpha_t(1 - \tau(\alpha_t^2 + \|v_t\|_2^2) + \tau) = (1 + \tau)\alpha_t - \tau\alpha_t^3 - \tau\alpha_t\|v_t\|_2^2; \\ v_{t+1} &= (1 - \tau(\alpha_t^2 + \|v_t\|_2^2))v_t + \tau M v_t. \end{aligned}$$

b)

$$\begin{aligned} (1 + \tau)\alpha_t - \tau\alpha_t^3 - \tau\alpha_t\|v_t\|_2^2 &= 1 + [(1 + \tau)\alpha_t - \tau\alpha_t^3 - \tau\alpha_t\|v_t\|_2^2 - 1] \\ &= 1 + [\alpha_t - 1 + \tau\alpha_t(1 - \alpha_t^2) - \tau\alpha_t\|v_t\|_2^2] \\ &= 1 + [(1 - \tau\alpha_t(\alpha_t + 1))(\alpha_t - 1) - \tau\alpha_t\|v_t\|_2^2]. \end{aligned}$$

4. We have

$$\begin{aligned} |1 - \alpha_t| &\leq \|x_t - u_1\|_2 \\ &\leq \|x_0 - u_1\|_2 \quad \text{from Eq. (4)} \\ &< \frac{1 - \lambda_2}{7}. \end{aligned}$$

Therefore, $1 - \left(\frac{1 - \lambda_2}{7}\right) < 1 - |1 - \alpha_t| \leq \alpha_t \leq 1 + |1 - \alpha_t| < 1 + \left(\frac{1 - \lambda_2}{7}\right)$.
And $\|v_t\| \leq \|x_t - u_1\|_2 < \frac{1 - \lambda_2}{7}$.

5. From Question 3.b),

$$|\alpha_{t+1} - 1| \leq |1 - \tau\alpha_t(\alpha_t + 1)| |\alpha_t - 1| + \tau|\alpha_t| \|v_t\|_2^2.$$

We prove the result by showing

$$|1 - \tau\alpha_t(\alpha_t + 1)| \leq 1 - \frac{5}{7}(1 - \lambda_2)\tau; \quad (8a)$$

$$\tau|\alpha_t| \|v_t\|_2^2 \leq \frac{8}{49}(1 - \lambda_2)\tau \|v_t\|_2. \quad (8b)$$

For Equation (8a), we must show

$$-\left(1 - \frac{5}{7}(1 - \lambda_2)\tau\right) \leq 1 - \tau\alpha_t(\alpha_t + 1) \leq 1 - \frac{5}{7}(1 - \lambda_2)\tau.$$

The left-hand side is equivalent to

$$\tau \left(\alpha_t(\alpha_t + 1) + \frac{5}{7}(1 - \lambda_2) \right) \leq 2,$$

which is true because $\tau \leq \frac{1}{2}$, $\alpha_t \leq 1 + \left(\frac{1-\lambda_2}{7}\right) \leq \frac{8}{7}$ and $1 - \lambda_2 \leq 1$, hence

$$\tau \left(\alpha_t(\alpha_t + 1) + \frac{5}{7}(1 - \lambda_2) \right) \leq \frac{1}{2} \left(\frac{8}{7} \times \frac{15}{7} + \frac{5}{7} \right) = \frac{155}{98} \leq 2.$$

The right-hand side is equivalent to

$$\alpha_t(\alpha_t + 1) \geq \frac{5}{7}(1 - \lambda_2),$$

which is true because $\alpha_t \geq 1 - \left(\frac{1-\lambda_2}{7}\right) \geq \frac{6}{7}$ and $\alpha_t + 1 \geq 1$, so

$$\alpha_t(\alpha_t + 1) \geq \frac{6}{7} \geq \frac{5}{7} \geq \frac{5}{7}(1 - \lambda_2).$$

Equation (8a) is proved.

For Equation (8b), we must show that

$$|\alpha_t| \|v_t\|_2 \leq \frac{8}{49}(1 - \lambda_2).$$

We have already said that $\alpha_t \leq \frac{8}{7}$, and we know from the previous question that $\|v_t\|_2 < \frac{1-\lambda_2}{7}$.

$$|\alpha_t| \|v_t\|_2 \leq \frac{8}{7} \times \frac{1 - \lambda_2}{7} = \frac{8}{49}(1 - \lambda_2).$$

6. a) From Question 3.a), $v_{t+1} = H_t v_t$, where

$$H_t = (1 - \tau(\alpha_t^2 + \|v_t\|_2^2)) \text{Id} + \tau M.$$

On the subspace $\text{Vect}\{u_2, \dots, u_d\}$, which v_t belongs to, M represents a symmetric linear operator with eigenvalues $\lambda_2, \dots, \lambda_d$. Therefore, H_t is a symmetric linear operator, with eigenvalues $(1 - \tau(\alpha_t^2 + \|v_t\|_2^2)) + \tau\lambda_k$ for $k = 2, \dots, d$.

All these eigenvalues are nonnegative,² hence the operator norm of H_t (still restricted to the subspace $\text{Vect}\{u_2, \dots, u_d\}$) is its largest eigenvalue:

$$(1 - \tau(\alpha_t^2 + \|v_t\|_2^2)) + \tau\lambda_2 = 1 - \tau(\alpha_t^2 + \|v_t\|_2^2 - \lambda_2),$$

which implies

$$\|v_{t+1}\|_2 \leq (1 - \tau(\alpha_t^2 + \|v_t\|_2^2 - \lambda_2)) \|v_t\|_2.$$

b) We have seen in the previous question (in footnote) that $0 \leq \tau\lambda_2 \leq 1 - \tau(\alpha_t^2 + \|v_t\|_2^2 - \lambda_2)$. Let us show that $1 - \tau(\alpha_t^2 + \|v_t\|_2^2 - \lambda_2) \leq 1 - \frac{5}{7}(1 - \lambda_2)\tau$, which is equivalent to

$$\alpha_t^2 + \|v_t\|_2^2 \geq \frac{5}{7} + \frac{2}{7}\lambda_2$$

We recall that $\alpha_t \geq 1 - \left(\frac{1-\lambda_2}{7}\right)$.

$$\begin{aligned} \alpha_t^2 + \|v_t\|_2^2 &\geq \left(1 - \left(\frac{1-\lambda_2}{7}\right)\right)^2 \\ &= 1 - 2\frac{1-\lambda_2}{7} + \left(\frac{1-\lambda_2}{7}\right)^2 \\ &= \frac{5}{7} + \frac{2}{7}\lambda_2 + \left(\frac{1-\lambda_2}{7}\right)^2 \\ &\geq \frac{5}{7} + \frac{2}{7}\lambda_2. \end{aligned}$$

²Observe that $\tau(\alpha_t^2 + \|v_t\|^2) \leq \frac{1}{2}\|x_t\|_2^2 \leq \frac{1}{2}(1 + \|x_t - u_1\|_2)^2 \leq \frac{1}{2}\left(\frac{8}{7}\right)^2 < 1$.

7. a) Combining the last questions, we get

$$\begin{aligned} |\alpha_{t+1} - 1| &\leq \left(1 - \frac{5}{7}(1 - \lambda_2)\tau\right) |\alpha_t - 1| + \frac{8}{49}(1 - \lambda_2)\tau \|v_t\|_2; \\ \|v_{t+1}\|_2 &\leq \left(1 - \frac{5}{7}(1 - \lambda_2)\tau\right) \|v_t\|_2. \end{aligned}$$

Expressed in terms of ℓ^2 -norms, this implies

$$\begin{aligned} &\|(|\alpha_{t+1} - 1|, \|v_{t+1}\|_2)\|_2 \\ &\leq \left\| \left(1 - \frac{5}{7}(1 - \lambda_2)\tau\right) (|\alpha_t - 1|, \|v_t\|_2) + \left(\frac{8}{49}(1 - \lambda_2)\tau \|v_t\|_2, 0\right) \right\|_2 \\ &\leq \left(1 - \frac{5}{7}(1 - \lambda_2)\tau\right) \|(|\alpha_t - 1|, \|v_t\|_2)\|_2 + \frac{8}{49}(1 - \lambda_2)\tau \|v_t\|_2 \\ &\quad \text{(triangular inequality)} \\ &\leq \left(1 - \frac{5}{7}(1 - \lambda_2)\tau + \frac{8}{49}(1 - \lambda_2)\tau\right) \|(|\alpha_t - 1|, \|v_t\|_2)\|_2 \\ &= \left(1 - \frac{27}{49}(1 - \lambda_2)\tau\right) \|(|\alpha_t - 1|, \|v_t\|_2)\|_2 \\ &\leq \left(1 - \left(\frac{1 - \lambda_2}{2}\right)\tau\right) \|(|\alpha_t - 1|, \|v_t\|_2)\|_2. \end{aligned}$$

b) We prove Inequality (4) by iteration over t . For $t = 0$, it is true. Now, if it is true for some t , the previous question implies

$$\begin{aligned} \|x_{t+1} - u_1\|_2 &= \sqrt{|\alpha_{t+1} - 1|^2 + \|v_{t+1}\|_2^2} \\ &\leq \left(1 - \frac{(1 - \lambda_2)\tau}{2}\right) \sqrt{|\alpha_t - 1|^2 + \|v_t\|_2^2} \\ &= \left(1 - \frac{(1 - \lambda_2)\tau}{2}\right) \|x_t - u_1\|_2 \\ &\leq \left(1 - \frac{(1 - \lambda_2)\tau}{2}\right)^{t+1} \|x_0 - u_1\|_2. \end{aligned}$$

Answer of Exercise 13

1. In the previous exercise, we have seen that, for all $x \in \mathbb{R}^d$,

$$\nabla f(x) = \|x\|_2^2 x - Mx.$$

We compute the Hessian of f using the Taylor expansion at order 1 : $\nabla f(x+h) = \nabla f(x) + \text{Hess}f(x)(h) + o(h)$, for all $x \in \mathbb{R}^d$ and $h \in \mathbb{R}^d$ going to zero. For any $x, h \in \mathbb{R}^d$,

$$\begin{aligned} \nabla f(x+h) &= (\|x+h\|_2^2)(x+h) - M(x+h) \\ &= (\|x\|^2 + 2\langle x, h \rangle + o(h))(x+h) - Mx - Mh \\ &= \nabla f(x) + \|x\|^2 h + 2\langle x, h \rangle x - Mh + o(h). \end{aligned}$$

Therefore, for any $x, h \in \mathbb{R}^d$,

$$\begin{aligned} \text{Hess}f(x)(h) &= \|x\|^2 h + 2\langle x, h \rangle x - Mh \\ &= \|x\|^2 h + 2xx^T h - Mh \\ &= (\|x\|^2 I_d + 2xx^T - M)h, \end{aligned}$$

which implies the desired result.

2. a) We first observe that, for any $x \in \mathbb{R}^d$, if x is first-order critical, then $\nabla f(x) = 0$, hence

$$Mx = \|x\|_2^2 x,$$

so that x is an eigenvector of M . We can therefore restrict our search for first-order critical points to the set of eigenvectors, which we will more conveniently write

$$\{rv, r \in \mathbb{R}^+, v \text{ is a unit eigenvector of } M\}.$$

Let v be a unit eigenvector, and r be in \mathbb{R}^+ . Let us denote λ the eigenvalue associated to v . When is rv a first-order critical point?

$$\begin{aligned} \nabla f(rv) = 0 &\iff \|rv\|^2(rv) - M(rv) = 0 \\ &\iff r^3v - r\lambda v = 0 \\ &\iff r(r^2 - \lambda) = 0. \end{aligned}$$

Therefore, rv is a first-order critical point if and only if $r = 0$ or³ $r = \sqrt{\lambda}$.

³As r is in \mathbb{R}^+ , it is impossible that $r = -\sqrt{\lambda}$.

The set of first-order critical points is

$$\{0\} \cup \{\sqrt{\lambda}v, v \text{ unit eigenvector of } M, \lambda \text{ associated eigenvalue}\}.$$

- b) A second-order critical point is a first-order critical point at which the Hessian is positive semidefinite.

First, we see that 0 is not a second-order critical point:

$$\text{Hess}f(0) = -M,$$

which is not semidefinite positive, since 1 is an eigenvalue of M , hence -1 is an eigenvalue of $-M$.

Now, let v be a unit eigenvector of f and λ its associated eigenvalue. Let us determine when $\sqrt{\lambda}v$ is a second-order critical point.

$$\text{Hess}f(\sqrt{\lambda}v) = \lambda I_d + 2\lambda vv^T - M.$$

If $\lambda = \lambda_1$, then this matrix is semidefinite positive, hence $\sqrt{\lambda}v$ is a second-order critical point. Indeed, for any $h \in \mathbb{R}^d$,

$$\begin{aligned} \langle h, \text{Hess}f(\sqrt{\lambda}v)h \rangle &= \lambda \|h\|^2 + 2\lambda \langle v, h \rangle^2 - \langle h, Mh \rangle \\ &\geq \lambda \|h\|^2 + 2\lambda \langle v, h \rangle^2 - \lambda_1 \|h\|^2 \\ &= 2\lambda \langle v, h \rangle^2 \\ &\geq 0. \end{aligned}$$

Conversely, if $\lambda \neq \lambda_1$, then the matrix is not semidefinite positive. Indeed, let h be an eigenvector associated with eigenvalue λ_1 . Then $\langle h, v \rangle = 0$ (since M is symmetric, the eigenvectors associated with different eigenvalues are orthonormal). Therefore,

$$\begin{aligned} \langle h, \text{Hess}f(\sqrt{\lambda}v)h \rangle &= \lambda \|h\|^2 + 2\lambda \langle v, h \rangle^2 - \langle h, Mh \rangle \\ &= \lambda \|h\|^2 - \langle h, Mh \rangle \\ &= \lambda \|h\|^2 - \lambda_1 \|h\|^2 \\ &< 0. \end{aligned}$$

Consequently, $\sqrt{\lambda}v$ is a second-order critical point if and only if $\lambda = \lambda_1 = 1$. Therefore, f has exactly two second-order critical points, which are u_1 and $-u_1$, where u_1 is a unit eigenvector associated with eigenvalue 1.

3. The map f is polynomial in the coordinates of x , therefore analytic. It is also coercive, as shown in Question 1.b) of the previous exercise. Therefore, we can apply a theorem seen in class (more precisely, the remark after the theorem: the theorem requires f to have a Lipschitz gradient, but the remark removes this assumption): for almost any x_0 , provided that the stepsize is small enough, gradient descent on f starting at x_0 converges towards a second-order critical point. Since the second-order critical points are the global minimizers (recall that we had computed the minimizers in Question 1.d) of the previous exercise), the result follows.

References

- Y. Chi, Y. M. Lu, and Y. Chen. Nonconvex optimization meets low-rank matrix factorization: An overview. *IEEE Transactions on Signal Processing*, 67(20):5239–5269, 2019.
- R. Gerchberg and W. Saxton. A practical algorithm for the determination of phase from image and diffraction plane pictures. *Optik*, 35(2):237–246, 1972.