Non-convex inverse problems: exercises

Irène Waldspurger

waldspurger@ceremade.dauphine.fr

January and February 2023

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 and partial uniqueness 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*.

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

This inequality tells us that the number $\frac{\lambda_1}{\lambda_d}$, which is called the *condition number* of A, controls the *stability* of the problem: if $\frac{\lambda_1}{\lambda_d}$ is close to 1, then a small error ϵ on y only causes a small error on x_*. If, on the other hand, $\frac{\lambda_1}{\lambda_d} \gg 1$, then x_ϵ can be very different from x_* even if ϵ is small.

Exercise 2: 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

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

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 $z, z' \in \mathbb{R}^d$,

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

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

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

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 $(||x_k - x_*||)_{k\in\mathbb{N}}$ 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 $x_* \in \cap_{s\leq S} C_s$,

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

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

Exercise 3: 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

$$\begin{aligned} &\text{find } x \in \mathbb{R}^d \\ &\text{such that } |\langle x, v_s\rangle| = y_s, \forall s \leq m, \end{aligned} \tag{Real-PR}$$

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 $S \subset \{1, \dots, m\}$,

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

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 $y_1, \dots, y_m$.

Exercise 4: 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

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

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

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

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
$$T_* = \{1, 2, \dots, \mathrm{Card}(T_*)\}$$
and that the coordinates of h are non-increasing, in absolute value, outside T_*:
$$|h_{\mathrm{Card}(T_*)+1}| \geq |h_{\mathrm{Card}(T_*)+2}| \geq ... \geq |h_d|.$$
Let us partition $\{\mathrm{Card}(T_*) + 1, \dots, d\}$ into sets $T_1, T_2, \dots, T_L$ of size $3k$:

$$T_1 = \{\mathrm{Card}(T_*) + 1, \dots, \mathrm{Card}(T_*) + 3k\},$$
$$T_2 = \{\mathrm{Card}(T_*) + 3k + 1, \dots, \mathrm{Card}(T_*) + 6k\},$$
$$\dots$$

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 5: guarantees for nuclear norm minimization

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

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

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:

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

We can assume that the λ_s are nonnegative and ordered: $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_{d_2} \geq 0$.

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

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

$$H = \begin{array}{c@{}l} \begin{array}{cc} \overset{r}{\longleftrightarrow} & \overset{d_2-r}{\longleftrightarrow} \end{array} & \\ \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} & \begin{array}{l} \updownarrow r \\ \updownarrow d_1 - r \end{array} \end{array}.$$

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}.)

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

We define matrices $H_{c,1}, \ldots, H_{c,L}$ 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

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

With this definition, $H_{c,1}, \ldots, H_{c,L}$ 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 $\mathrm{rank}(H_0) \leq 2r$ 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 6: 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 $\tau_1, \dots, \tau_S$ are distinct elements of $[0;1[$. We assume that we have access to its $2S$ lowest-frequency Fourier coefficients:

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

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

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

Express it as a function of $\tau_1, \dots, \tau_S$.

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

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

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

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

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

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

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

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 7: super-resolution via semidefinite programming

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

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

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

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

We have seen 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 $P\left(e^{2\pi it}\right) = \sum_{k=-2N}^{2N} p_k e^{2\pi ikt}$ 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, $P\left(e^{2\pi it}\right) \in \mathbb{R}^+$ 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

$$P\left(e^{2\pi it}\right) = \sum_{k=1}^{n} \left|Q_k\left(e^{2\pi it}\right)\right|^2.$$

1. Let $z \in \mathbb{C}^{2N+1}$ be any vector. Show that $\left|\sum_{k=-N}^{N} z_k e^{2\pi ikt}\right| \leq 1$ 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

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

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:

$$P_l\left(e^{2\pi it}\right) = \sum_{k=-N}^{N} p_k^{(l)} e^{2\pi ikt}.$$

   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 $d = -2N, \dots, 2N$,

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

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

$$\left(\begin{array}{ccc|c} & & & z_{-N} \\ & A & & \vdots \\ & & & z_N \\ \hline \overline{z_{-N}} & \dots & \overline{z_N} & 1 \end{array}\right) \succeq 0.$$

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

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

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

Exercise 8: Fejér-Riesz theorem

In this exercise, we prove Fejér-Riesz' theorem, stated in the previous exercise. Let $N \in \mathbb{N}^*$ be fixed.

1. Let P be a trigonometric polynomial with degree at most $2N$. We assume it can be written as the sum of the squared modulus of trigonometric polynomials $Q_1, \dots, Q_n$:

$$P\left(e^{2\pi it}\right) = \sum_{k=1}^{n} \left|Q_k\left(e^{2\pi it}\right)\right|^2.$$

Show that, for all $t \in \mathbb{R}$, $P\left(e^{2\pi it}\right)$ belongs to $\mathbb{R}^+$.

2. Conversely, let $P\left(e^{2\pi it}\right) = \sum_{k=-2N}^{2N} p_k e^{2\pi ikt}$ be a trigonometric polynomial with degree at most $2N$, such that $P\left(e^{2\pi it}\right) \in \mathbb{R}^+$ for all $t \in \mathbb{R}$.
We assume $p_{2N} \neq 0$.[1]
Let $\tilde{P}(X) = X^{2N} \sum_{k=-2N}^{2N} p_k X^k$ be the "standard" polynomial associated to P. It has degree $4N$.
   a) Let $z_1, \dots, z_{4N}$ be the roots of $\tilde{P}$ in $\mathbb{C}$ (counted with multiplicity). Express P as a function of $z_1, \dots, z_{4N}$ and p_{2N}.
   b) Show that, for any $z \in \mathbb{C}$,
   $$\tilde{P}(z) = z^{4N} \overline{\tilde{P}\left(\frac{1}{\bar{z}}\right)}.$$
   [Hint: use the fact that $P\left(e^{2\pi it}\right)$ is a real number, for any $t \in \mathbb{R}$.]
   c) Deduce from the previous question that, for any $z \in \mathbb{C}$, if z is a root of $\tilde{P}$, then $\frac{1}{\bar{z}}$ is also a root of $\tilde{P}$, with the same multiplicity.
   d) Show that, for any $z \in \mathbb{C}$ such that $|z| = 1$, if z is a root of $\tilde{P}$, its multiplicity is even.
   [Hint: show that, if the multiplicity is odd, the sign of P changes in the neighborhood of $e^{2\pi it} \stackrel{def}{=} z$.]
   e) Show that there exists a trigonometric polynomial Q with degree N such that
   $$P(e^{2\pi it}) = \left|Q\left(e^{2\pi it}\right)\right|^2.$$
   [Remark: this establishes the second property of Fejér-Riesz' theorem with $n = 1$.]

Exercise 9: alternating projections for phase retrieval

We consider a generic phase retrieval problem:

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

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

[1] This is without loss of generality. If $p_{2N} = 0$, the same reasoning is true; it essentially suffices to replace $2N$ with the index of the smallest integer D for which $p_D \neq 0$.

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

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

   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 $\text{proj}_{\mathcal{E}} : \mathbb{C}^m \to \mathbb{C}^m$ such that,

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

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

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

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

Exercise 10: 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

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

(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, $f(x) = \frac{1}{4}\left(||x||_2^4 - 2\langle x, Mx\rangle + ||M||_F^2\right)$.
   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 $t \in \mathbb{N}$,

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

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

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

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,

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

   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 $1 - \left(\frac{1-\lambda_2}{7}\right) \leq \alpha_t \leq 1 + \left(\frac{1-\lambda_2}{7}\right)$ 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 11: rank 1 approximation (global convergence)

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

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

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 $y \in \mathrm{Range}(A)$. This solution, which we denote x_*, is unique if the set
$$\{x \in \mathbb{R}^d \text{ such that } Ax = Ax_*\} = \{x_* + h, h \in \mathrm{Ker}(A)\}$$
is the singleton $\{x_*\}$. This happens if and only if A is injective (that is $\mathrm{Ker}(A) = \{0\}$).

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
$$\{v \in \mathrm{Vect}\{v_1, \dots, v_{k-1}\}^\perp, |||v||_2 = 1\}$$
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 $v_{k'}$,
$$v_{k'} \in \mathrm{Vect}\{v_1, \dots, v_{k'-1}\}^\perp \subset \mathrm{Vect}\{v_k\}^\perp \quad \Rightarrow \quad \langle v_{k'}, v_k \rangle = 0.$$
As a consequence, for any $\theta \in \mathbb{R}$,
$$|| \cos(\theta) v_k + \sin(\theta) v_{k'} ||_2 = \sqrt{\cos^2(\theta) ||v_k||_2^2 + \sin^2(\theta) ||v_{k'}||_2^2} = 1. \quad (5)$$
In addition, v_k is in $\mathrm{Vect}\{v_1, \dots, v_{k-1}\}^\perp$ and $v_{k'}$ is in $\mathrm{Vect}\{v_1, \dots, v_{k'-1}\}^\perp \subset \mathrm{Vect}\{v_1, \dots, v_{k-1}\}^\perp$, so
$$\cos(\theta) v_k + \sin(\theta) v_{k'} \in \mathrm{Vect}\{v_1, \dots, v_{k-1}\}^\perp. \quad (6)$$

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

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

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

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

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

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 $\mathrm{Vect}\{v_1, \dots, v_{k-1}\}^{\perp}$, and $||v_{k'}||_2 = 1$. Hence, from the definition of v_k,

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

d) Let D be the smallest index such that $\lambda_D = 0$ (it is possible that $\lambda_k \neq 0$ for all $k \le d$, 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 complete it to an orthonormal basis $(u_1, \dots, u_m)$ of $\mathbb{R}^m$, which defines $u_D, \dots, u_m$.
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 $V^T V = VV^T = \mathrm{Id}$, 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

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

We respectively denote $(x_{V,k})_{k \le d}$, $(x_{V,k}^{(\epsilon)})_{k \le d}$, $(y_{U,k})_{k \le m}$ and $(\epsilon_{U,k})_{k \le m}$ the coordinates of Vx_*, Vx_ϵ, $U^T y$ and $U^T \epsilon$. From Equation (7), for all $k \le d$,

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

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

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

Therefore,

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

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,

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

Answer of Exercise 4

1. The vector x_* is feasible for the problem (Basis Pursuit): $A x_* = y$. 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,

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

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 $s \in T_l$,

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

From this, we deduce that

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

b)

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

$$\leq \frac{1}{\sqrt{3k}} \sum_{l=1}^{L} ||h_{T_l}||_1$$

$$= \frac{||h_{T_*^c}||_1}{\sqrt{3k}}$$

$$\leq \frac{||h_{T_*}||_1}{\sqrt{3k}} \quad \text{from the first question.}$$

c) By Cauchy-Schwarz,

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

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 $h = h_{T_* \cup T_1} + h_{T_2} + \cdots + h_{T_L}$, we have

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

The vector $h_{T_* \cup T_1}$ has at most $\mathrm{Card}(T_*) + \mathrm{Card}(T_1) \leq k + 3k = 4k$ 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 $l \in \{2, \ldots, L\}$,

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

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

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

Using also Question 2.c,

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

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

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

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.

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

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

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

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

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

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

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

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

$= 0$ otherwise.

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:

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

   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 $BB^* = A - zz^*$).[2] 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 $A = zz^* + \sum_{l=1}^n p^{(l)} p^{(l)*}$.

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 $G \succeq 0$.

$$\begin{aligned}
(G \succeq 0) \iff & (\forall h \in \mathbb{C}^{2N+2}, h^* G h \geq 0) \\
\iff & \left(\forall \tilde{h} \in \mathbb{C}^{2N+1}, u \in \mathbb{C}, \left(\begin{smallmatrix} \tilde{h} \\ u \end{smallmatrix}\right)^* G \left(\begin{smallmatrix} \tilde{h} \\ u \end{smallmatrix}\right)) \geq 0\right) \\
\iff & \left(\forall \tilde{h} \in \mathbb{C}^{2N+1}, u \in \mathbb{C}, \tilde{h}^* A \tilde{h} + 2\mathrm{Re}\left(u \left\langle \tilde{h}, z \right\rangle\right) + |u|^2 \geq 0\right) \\
\iff & \Big(\forall \tilde{h} \in \mathbb{C}^{2N+1}, t \in \mathbb{R}, \phi \in \mathbb{R}, \\
& \qquad \tilde{h}^* A \tilde{h} + 2\mathrm{Re}\left(t e^{i\phi} \left\langle \tilde{h}, z \right\rangle\right) + |t e^{i\phi}|^2 \geq 0\Big) \\
\iff & \Big(\forall \tilde{h} \in \mathbb{C}^{2N+1}, t \in \mathbb{R}, \phi \in \mathbb{R}, \\
& \qquad \tilde{h}^* A \tilde{h} + 2t\mathrm{Re}\left(e^{i\phi} \left\langle \tilde{h}, z \right\rangle\right) + t^2 \geq 0\Big)
\end{aligned}$$

[2] All semidefinite positive matrices have square roots; it is most easily proved when writing the semidefinite matrix in an eigenvector basis.

$$\overset{(a)}{\iff} \left(\forall \tilde{h} \in \mathbb{C}^{2N+1}, \phi \in \mathbb{R}, \tilde{h}^* A \tilde{h} - \left(\mathrm{Re}\left(e^{i\phi}\left\langle \tilde{h}, z \right\rangle\right)\right)^2 \geq 0\right)$$
$$\overset{(b)}{\iff} \left(\forall \tilde{h} \in \mathbb{C}^{2N+1}, \tilde{h}^* A \tilde{h} - |\left\langle \tilde{h}, z \right\rangle|^2 \geq 0\right)$$
$$\iff \left(\forall \tilde{h} \in \mathbb{C}^{2N+1}, \tilde{h}^* (A - zz^*) \tilde{h} \geq 0\right)$$
$$\iff A - zz^* \succeq 0.$$

Equivalence (a) is true because, for any $\tilde{h}$ and ϕ, it holds that the polynomial $t \to \tilde{h}^* A \tilde{h} + 2t\mathrm{Re}\left(e^{i\phi}\left\langle \tilde{h}, z \right\rangle\right) + t^2$ is nonnegative over $\mathbb{R}$ if and only if its discriminant is nonpositive:

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

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

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

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

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

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

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

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

We denote $p^{(1)}, \dots, p^{(n)}$ the vectors of their coefficients and set $A = zz^* + \sum_{l=1}^{n} p^{(l)}p^{(l)*}$. 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 $\left|\sum_{k=-N}^{N} z_k e^{2\pi ikt}\right| \leq 1$ 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)

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

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

Answer of Exercise 10

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

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

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

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

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

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

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

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

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

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

As a consequence, $M x_{min} = ||x_{min}||_2^2 x_{min}$, 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, $||x_{min}|| = \sqrt{\lambda_k}$.

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 $k \le d$ and v as above,

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

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)

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

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

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

b)

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

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

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

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

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 $\alpha_t \geq 1 - \left(\frac{1-\lambda_2}{7}\right) \geq \frac{6}{7}$, so

$$\alpha_t(\alpha_t + 1) \geq \frac{6}{7} \times \frac{13}{7} \geq \frac{5}{7} \geq \frac{5}{7}(1 - \lambda_2).$$

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

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

On the subspace $\mathrm{Vect}\{u_2, \ldots, u_d\}$, which v_t belongs to, M represents a symmetric linear operator with eigenvalues $\lambda_2, \ldots, \lambda_d$. Therefore, H_t is a symmetric linear operator, with eigenvalues $(1 - \tau(\alpha_t^2 + ||v_t||_2^2)) + \tau\lambda_k$ for $k = 2, \ldots, d$.
All these eigenvalues are nonnegative,[3] hence the operator norm of H_t (still restricted to the subspace $\mathrm{Vect}\{u_2, \ldots, u_d\}$) is its largest eigenvalue:

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

which implies

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

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

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

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

7. a) Combining the last questions, we get

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

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

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

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

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

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.