Non-convex inverse problems: exam

March 20, 2023

1h30

You can use any written or printed material.

The subject is long, but you do not need to answer every question. Correctly answering all questions until 2.a) in Exercise 2 should be enough to get the maximal grade.

Exercice 1

Let $n \in \mathbb{N}^*$ be fixed, and let Sym_n denote the set of symmetric $n \times n$ matrices. Given $M \in \mathrm{Sym}_n$, we assume that we want to find its smallest eigenvalue by solving the following problem:

$$\begin{aligned} \text{minimize } & \langle Mx, x\rangle \\ \text{over all } & x \in \mathbb{R}^n \\ \text{such that } & ||x||_2 = 1. \end{aligned} \tag{Min Eig}$$

To simplify notation, we assume that M is diagonal:

$$M = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix},$$

where $\lambda_1 \le \lambda_2 \le \dots \le \lambda_n$ are the ordered eigenvalues.

1. Show that the optimal value of Problem (Min Eig) is λ_1, and that the minimizers are exactly the eigenvectors of M with norm 1, associated to eigenvalue λ_1.
2. Why can we call Problem (Min Eig) *non-convex*?
   [Be careful: there are two reasons.]
3. Show that Problem (Min Eig) is equivalent to the following low-rank matrix recovery problems (that is, the two problems have the same optimal value, and the minimizers of (Low Rank) are exactly the matrices of the form $x_0 x_0^* \in \mathbb{R}^{n\times n}$ for $x_0 \in \mathbb{R}^n$ a minimizer of (Min Eig)):

$$\begin{aligned} \text{minimize } & \mathrm{Tr}(MX) \\ \text{over all } & X \in \mathrm{Sym}_n, \\ \text{such that } & \mathrm{Tr}(X) = 1, \\ & X \succeq 0, \\ & \mathrm{rank}(X) = 1. \end{aligned} \tag{Low Rank}$$

4. We drop the rank 1 constraint to get the following convex relaxation:

$$\begin{aligned} \text{minimize } & \operatorname{Tr}(MX) \\ \text{over all } & X \in \operatorname{Sym}_n, \\ \text{such that } & \operatorname{Tr}(X) = 1, \\ & X \succeq 0. \end{aligned} \tag{Relaxation}$$

   a) Why haven't we added a nuclear norm somewhere?
   b) Show that (Low Rank) and (Relaxation) have the same optimal value and that minimizers of (Low Rank) are also minimizers of (Relaxation).

5. Let $x_0 \in \mathbb{R}^n$ be a solution of (Min Eig). In this question, we construct a dual certificate to show that $x_0 x_0^*$ is a solution of (Relaxation).[1]
   Let us introduce the Lagrangian function

$$\begin{aligned} \mathcal{L}: \quad \operatorname{Sym}_n \times \operatorname{Sym}_n \times \mathbb{R} & \to \mathbb{R} \\ (X, H, a) & \to \operatorname{Tr}(MX) - \operatorname{Tr}(HX) + a(\operatorname{Tr}(X) - 1). \end{aligned}$$

   a) Show that, for any $X \in \operatorname{Sym}_n$ which is feasible for Problem (Relaxation), it holds

$$\operatorname{Tr}(MX) \geq \sup_{\substack{H \in \operatorname{Sym}_n, a \in \mathbb{R} \\ \text{s.t. } H \succeq 0}} \mathcal{L}(X, H, a).$$

   b) Show that for any $H \in \operatorname{Sym}_n, a \in \mathbb{R}$,

$$\begin{aligned} \inf_{X \in \operatorname{Sym}_n} \mathcal{L}(X, H, a) &= -a \text{ if } M - H + aI_n = 0 \\ &= -\infty \text{ otherwise.} \end{aligned}$$

   [Hint: observe that, for all a, X, $a\operatorname{Tr}(X) = \operatorname{Tr}(aI_n X)$.]
   c) From the previous question, we define the dual of (Relaxation) as

$$\begin{aligned} \text{maximize } & -a \\ \text{over all } & a \in \mathbb{R}, H \in \operatorname{Sym}_n, \\ \text{such that } & H = M + aI_n, \\ & H \succeq 0. \end{aligned} \tag{Dual}$$

   Show that $H_0 = M - \lambda_1 I_n, a_0 = -\lambda_1$ is a feasible pair for Problem (Dual), and that

$$-a_0 = \operatorname{Tr}(M x_0 x_0^*).$$

   d) Deduce from the previous question that $x_0 x_0^*$ is a solution of (Relaxation).

[1] From the previous question, we already know that $x_0 x_0^*$ is a solution of (Relaxation); the point here is to give a different proof.

Exercice 2

We consider a super-resolution problem, where we want to recover an unknown measure $\mu_\star$ over $[0;1[$, which is the sum of a few diracs, from its low-frequency Fourier coefficients:

$$\hat{\mu}_\star[k] = \int_0^1 e^{-2\pi i k t} d\mu_\star(t), \quad k = -N, \dots, N.$$

Here, N is a (strictly) positive integer.
We assume that we know that

- $\mu_\star$ is the sum of only one dirac;
- the coefficient of this dirac is a real number.

We define the objective function

$$\begin{array}{rccc} F: & \mathbb{R}^2 & \to & \mathbb{R} \\ & (\alpha,\theta) & \to & \sum_{k=-N}^{N} |\widehat{(\alpha\delta_\theta)}[k] - \hat{\mu}_\star[k]|^2 \\ & & & = \sum_{k=-N}^{N} |\alpha e^{-2\pi i k\theta} - \hat{\mu}_\star[k]|^2. \end{array}$$

To simplify the exercise, we assume that $\mu_\star = \delta_0$, hence $\hat{\mu}_\star[k] = 1$ for all k.

1. Show that (α, θ) is a global minimizer of F if and only if $\alpha = 1$ and $\theta \in \mathbb{Z}$.
2. We imagine that we run gradient descent on F, with a stepsize $\tau \in \left]0; \frac{1}{8(2N+1)^3}\right[$, starting at some (α_0, θ_0). This defines a sequence of iterates $(\alpha_t, \theta_t)_{t\in\mathbb{N}}$.
   In this question, we show that, if $(\alpha_0, \theta_0) \in]0;2[\times \left]-\frac{1}{2N+1}; \frac{1}{2N+1}\right[$, then

   $$(\alpha_t, \theta_t) \overset{t\to+\infty}{\longrightarrow} (1, 0).$$

   a) What is the name of this type of results?
   b) For any θ, we define $D_N(\theta) = \sum_{k=-N}^{N} \cos(2\pi k\theta)$. Check that

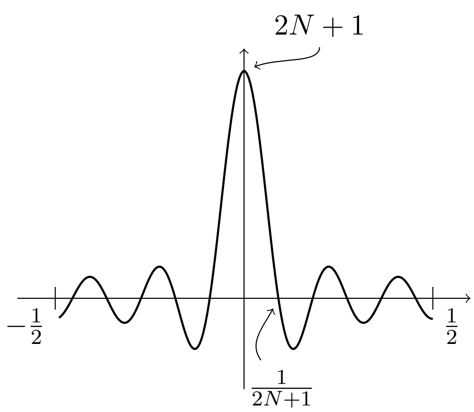
   $$F(\alpha,\theta) = (2N+1)\alpha^2 - 2\alpha D_N(\theta) + (2N+1), \quad \forall(\alpha,\theta) \in \mathbb{R}^2.$$

   c) Check that, for all $(\alpha, \theta) \in \mathbb{R}^2$,

   $$\nabla F(\alpha,\theta) = 2\begin{pmatrix} (2N+1)\alpha - D_N(\theta) \\ -\alpha D_N'(\theta) \end{pmatrix}.$$

Here are a few properties of D_N that you can use with no proof in the remaining questions.

- $D_N(\theta) > 0$ for all $\theta \in \left]-\frac{1}{2N+1}; \frac{1}{2N+1}\right[$.
- $D'_N(0) = 0$.
- $D'_N(\theta) < 0$ for all $\theta \in \left]0; \frac{1}{2N+1}\right[$.
- $D'_N(\theta) > 0$ for all $\theta \in \left]-\frac{1}{2N+1}; 0\right[$.
- $|D'_N(\theta)| \leq 4(2N+1)^3|\theta|$ for all $\theta \in \left]-\frac{1}{2N+1}; \frac{1}{2N+1}\right[$.



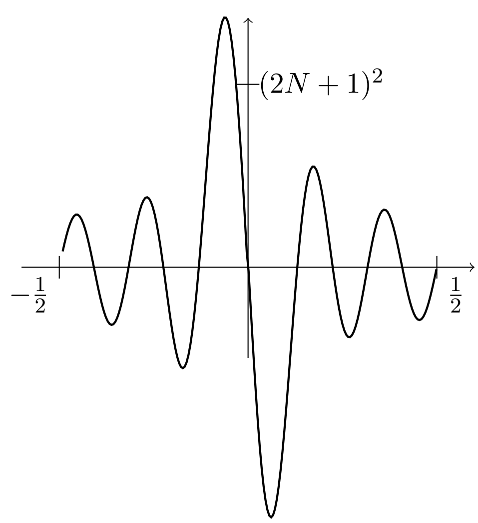


Figure 1: On the left, D_N ; on the right, D'_N.

d) For any $t \in \mathbb{N}$, show that, if $(\alpha_t, \theta_t) \in]0; 2[\times \left]-\frac{1}{2N+1}; \frac{1}{2N+1}\right[$, then

$$|\theta_{t+1}| \leq |\theta_t|.$$

e) For any $t \in \mathbb{N}$, show that, if (α_t, θ_t) belongs to $]0; 2[\times \left]-\frac{1}{2N+1}; \frac{1}{2N+1}\right[$, then $(\alpha_{t+1}, \theta_{t+1})$ also does.

f) (Difficult) Show that, if $(\alpha_0, \theta_0) \in]0; 2[\times \left]-\frac{1}{2N+1}; \frac{1}{2N+1}\right[$, then

$$(\alpha_t, \theta_t) \to (1, 0) \quad \text{when } t \to +\infty.$$