Convex analysis and optimisation

2020, November 18
9h30 – 12h

You may use the courses notes that were given to you and your handwritten notes during the exam, but you shouldn't use any online resource nor communicate with each other. A fraud could lead to a ban to pass university exams in France for up to five years.

The exam has two parts: Part I corresponds to the algorithmic part by J.C. Pesquet and Part II to the convex analysis part by Q. Mérigot. They are graded independently and account each for half of the final grade.

Part I: Article study

Answer the following questions concerning the article: "A Variational Inequality Model for the Construction of Signals from Inconsistent Nonlinear Equations" by P. L. Combettes and Z. C. Woodstock [1]. *All the provided answers should be properly justified.*

1. Is the purpose of this paper simply to solve Problem 1.1 ?
2. In which part of the article Lemma 2.5 is used ?
3. What is the expression of function S employed in (3.14) ? How can this expression be deduced from previous formulas ?
4. What is a weakly-convex function ?
5. Give an example of a weakly-convex function which is not convex.
6. Under which condition the set of solutions to Problem 1.1 is equal to the set of solutions to Problem 1.3 ?
7. What is the interest of using index set I_n instead of I at iteration n of Algorithm (4.15) ?
8. In Section 5.2, what is the use of the prescription modeled by F_2 and L_2 ? How would you suggest to adjust parameter ρ_2 to recover a less noisy signal in Figure 5.2(c).
9. Comment Figure 5.6.
10. Let $\mathcal{H}$ be a real Hilbert space. Let f, g, and h be functions in $\Gamma_0(\mathcal{H})$. In addition, suppose that h is Gâteaux differentiable and has a β-Lipschitzian gradient with $\beta \in \,]0, +\infty[$. We assume that there exists a common solution $\widehat{x} \in \mathcal{H}$ to the following two optimization problems:

$$\begin{cases} \underset{x\in\mathcal{H}}{\text{minimize}} \;\; f(x) + h(x) \\ \underset{x\in\mathcal{H}}{\text{minimize}} \;\; g(x). \end{cases}$$

How can we apply the framework in this article to find a common solution to these two problems ?

[1] https://arxiv.org/abs/2105.07380

Part II

The two exercises are independent. If you do not find the solution to a question, you can admit the answer and skip it. Given the length of this part, it will be possible to get the maximal grade without answering all the questions.

Exercise 1. Let $f_1, \dots, f_k \in \mathcal{C}^1(\mathbb{R}^n)$ be convex functions and let $f(x) = \max_{1 \le i \le k} f_i(x)$. We fix a point $x \in \mathbb{R}^n$ and our goal is to establish and use the subdifferential calculus rule

$$\partial f(x) = \operatorname{conv}\left(\{\nabla f_i(x)\}_{i \in I(x)}\right) = \left\{ \sum_{i \in I(x)} \lambda_i \nabla f_i(x) \mid \lambda \in \mathbb{R}_+^{I(x)} \text{ and } \sum_i \lambda_i = 1 \right\},$$

where $I(x) = \{i \in \{1, \dots, k\} \mid f(x) = f_i(x)\}$.

1. Prove the inclusion $\supseteq$.

2. We fix $z \in \mathbb{R}^n$, and we consider for now the following optimization problem

$$P_z = \inf_{(y,r) \in \mathbb{R}^n \times \mathbb{R} \text{ s.t. } \forall i, f_i(y) \le r} r - \langle y | z \rangle.$$

Prove that $(y, r) \in \mathbb{R}^n \times \mathbb{R}$ minimizes P_z if and only $r = f(y)$ and if there exists $\lambda \in \mathbb{R}_+^{I(y)}$ such that

$$\begin{cases} \sum_{i \in I(y)} \lambda_i = 1, \\ z = \sum_{i \in I(y)} \lambda_i \nabla f_i(y). \end{cases}$$

3. Prove that $f^*(z) = -P_z$, and show that y is a minimizer of P_z if and only if $z \in \partial f(y)$ if and only if $z \in \operatorname{conv}\left(\{\nabla f_i(y)\}_{i \in I(y)}\right)$.

4. *Application.* Let $y_1, \dots, y_k \in \mathbb{R}^n$ be distinct and let $f_i(x) = \frac{1}{2}\|x - y_i\|^2$.

(i) Prove that the function $f = \max_i f_i$ admits a unique minimizer on $\mathbb{R}^n$.

(ii) Compute $\partial f(x)$ for any $x \in \mathbb{R}^d$.

(iii) Deduce a characterization of the unique minimizer of f.

Exercise 2. We consider an optimization problem over the space of square n-by-n symmetric matrices $\mathcal{X} := \mathcal{S}_n$, endowed with the scalar product

$$\langle A | B \rangle_{\mathcal{X}} = \operatorname{Tr}(A^T B) = \sum_{1 \le i,j \le n} A_{ij} B_{ij}$$

and the Euclidean norm $\|A\|_{\mathcal{X}} = |\langle A | A \rangle_{\mathcal{X}}|^{1/2}$. The set of positive definite matrices is denoted $\mathcal{S}_n^{++} \subseteq \mathcal{X}$, and we set $\mathcal{S}_n^{--} = -\mathcal{S}_n^{++}$. We denote $\det(A)$ the determinant of a symmetric matrix $A \in \mathcal{X}$, $\operatorname{Tr}(A)$ its trace, and $\lambda_1(A) \le \dots \le \lambda_n(A)$ its eigenvalues. Then,

$$\det(A) = \prod_i \lambda_i(A), \quad \operatorname{Tr}(A) = \sum_i \lambda_i(A).$$

A (centered) *ellipsoid* is a set of the form

$$\mathcal{E}(A) = \{x \in \mathbb{R}^n \mid \langle x | Ax \rangle \le 1\}$$

where A belongs to $\mathcal{S}_n^{++}$. Finally we consider k points $y_1, \dots, y_k$ in $\mathbb{R}^n$, and we consider the problem of finding the ellipsoid $\mathcal{E}(A)$ containing those points with minimal volume.

1. Briefly justify that the problem can be formulated as

$$P := \inf_K F(A), \text{ where } F(A) = \begin{cases} -\ln\det A & \text{if } A \in \mathcal{S}_n^{++} \\ +\infty & \text{if not} \end{cases} \tag{1}$$
$$\text{and } K = \{A \in \mathcal{X} \mid \forall i \in \{1,\dots,k\}, \langle y_i | A y_i\rangle \leq 1\}.$$

2. Prove that $\det(\mathrm{Id} + tV) = 1 + t\,\mathrm{Tr}(V) + o(t)$. Deduce that $\nabla\det(\mathrm{Id}) = \mathrm{Id}$ and that if $A \in \mathcal{S}_n^{++}$, then $\nabla F(A) = -A^{-1}$.

3. The goal of this question is to prove that F is convex and lsc.

a. Prove that for any $B \in \mathcal{S}_n^{--} = -\mathcal{S}_n^{++}$, $F^*(B) = F(-B) - n$.

b. Prove that for any $B \in \mathcal{S}_n \setminus \mathcal{S}_n^{--}$, $F^*(B) = +\infty$.
(Hint: prove first that there exists $A \in \mathcal{S}_n^{++}$ such that $\langle A|B\rangle_{\mathcal{X}} \geq 0$)

c. Conclude.

4. Compute the subdifferential of $\partial i_K(A)$ for $A \in K$.
(Hint: you may let $g_i(A) = \langle y_i|Ay_i\rangle - 1$ and $I(A) = \{i \in \{1,\dots,k\} \mid g_i(A) = 0\}$.)

5. Deduce a necessary and sufficient optimality condition for (1).

6. Prove that if $y_1, \dots, y_n$ span the whole space, then (1) has a unique minimizer.