Optimization – Exercises

Day 1

Let $(H, \langle \cdot, \cdot \rangle)$ be a real Hilbert space. We denote $\| \cdot \|$ the norm derived by the scalar product.

Exercise 1 (Necessary and sufficient optimality conditions).
Let $f \colon H \longrightarrow \mathbb{R}$ be a twice differentiable function. Show that if x is a local minimizer of f, then

$$\nabla f(x) = 0$$
$$\nabla^2 f(x) \geq 0$$

Is the first order condition a sufficient condition for x to be a local minimizer ? If no, give an example. What assumption can you make for this condition to be an equivalence ?

Exercise 2 (Caracterizations of convex functions).
Let $f \colon H \longrightarrow \mathbb{R}$ be a twice differentiable function. Show the following equivalences :

1. f is convex if, and only if,
$$\forall (x, y) \in H \times H,\ f(y) \geqslant f(x) + \langle \nabla f(x) \mid y - x \rangle.$$
2. f is convex if, and only if,
$$\forall x \in H,\ \nabla^2 f(x) \geq 0,$$

where $\nabla^2 f(x)$ is the hessian of f at x.

Exercise 3 (Squared distance function).
Let A be a nonempty closed convex subset of H. We consider the function "squared distance to A" defined for all $x \in H$ by
$$g(x) = \inf_{y \in A} \|x - y\|^2.$$

1. Show that g is convex.
2. Show that g is Fréchet differentiable, with $\nabla g(x) = 2(x - \mathrm{p}_A(x))$, where p_A denotes the projection on A.

Exercise 4 (Minimization of a quadratic function).
Let $A \in \mathcal{S}_n^{++}(\mathbb{R})$ (set of symmetric positive definite matrices of $\mathbb{R}^{n \times n}$) and $b \in \mathbb{R}^n$. Let f be defined for all $x \in \mathbb{R}^n$ by
$$f(x) = \frac{1}{2} \langle Ax, x \rangle - \langle b, x \rangle.$$
Show that f admits a unique minimizer and give an expression of this minimizer.

Exercise 5 (Convex optimization exam 2019).
Let $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ be a convex, differentiable and bounded function on $\mathbb{R}^n$. Show f is constant.

Exercise 6 (About ε-minimizers).
Let $f \colon \mathbb{R}^n \longrightarrow \mathbb{R}$ be a continuous function bounded from below on $\mathbb{R}^n$. Let $\varepsilon > 0$ and u a ε-minimizer of f, i.e. u satisfies
$$f(u) \leqslant \inf_{x \in \mathbb{R}^n} f(x) + \varepsilon.$$

Let $\lambda > 0$ and consider

$$g\colon x \in \mathbb{R}^n \mapsto g(x) := f(x) + \frac{\varepsilon}{\lambda}\|x - u\|.$$

1. Show there exists $v \in \mathbb{R}^n$ which minimizes g on $\mathbb{R}^n$. Show this point v satisfies the following conditions :
   (i) $f(v) \leqslant f(u)$,
   (ii) $\|u - v\| \leqslant \lambda$,
   (iii) $\forall x \in \mathbb{R}^n,\ f(v) \leqslant f(x) + \frac{\varepsilon}{\lambda}\|x - v\|$.
2. Suppose in addition that f is differentiable on $\mathbb{R}^n$. Show that for all $\epsilon > 0$, there exists $x_\epsilon \in \mathbb{R}^n$ such that
$$\|\nabla f(x_\epsilon)\| \leqslant \epsilon.$$

Exercise 7.
Let $\mathcal{O} = \mathcal{S}_n^{++}(\mathbb{R})$ be the (open) set of symmetric positive definite matrices of $\mathbb{R}^{n\times n}$. $\mathcal{O}$ is endowed with the scalar product $\langle U, V\rangle = \mathrm{Tr}(UV)$. Let $A \in \mathcal{O}$ and f be defined for all $X \in \mathcal{O}$ by

$$f(X) = \mathrm{Tr}(X^{-1}) + \mathrm{Tr}(AX).$$

1. Show there exists a minimizer to f on $\mathcal{O}$. *Hint : you may use the inequality* $\mathrm{Tr}(UV) \geqslant \sum_{i=1}^n \lambda_i(U)\lambda_{n-i+1}(V)$, *where all eigenvalues* $\lambda_1, \ldots, \lambda_n$ *are in descending order ; i.e.,* $\lambda_1 \geqslant \cdots \geqslant \lambda_n$.
2. Find the minimizer and the optimal value of f.

Exercise 8 (Penalty method).
Let $F\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ be a lower semi-continuous function, coercive on $\mathbb{R}^n$. Let C be a closed set of $\mathbb{R}^n$ with $\mathrm{dom}(f) \cap C \neq \varnothing$. We seek to solve the constrained problem

$$\begin{aligned} \underset{x\in\mathbb{R}^n}{\text{minimize}} \quad & F(x) \\ \text{s.t.} \quad & x \in C. \end{aligned} \tag{$\mathcal{P}$}$$

Let $R\colon \mathbb{R}^n \longrightarrow \mathbb{R}^+$ be a lower semi-continuous function such that

$$R(x) = 0 \quad \Longleftrightarrow \quad x \in C.$$

R is called penalty function as it assigns a positive cost to any point that is not in the constraint set C. Let $(\gamma_k)_{k\in\mathbb{N}}$ be a nondecreasing sequence of positive reals satisfying $\lim_{k\to+\infty} \gamma_k = +\infty$. We denote by $(\mathcal{P}_k)$ the following penalized problem :

$$\underset{x\in\mathbb{R}^n}{\text{minimize}} \quad F_{\gamma_k}(x) := F(x) + \gamma_k R(x). \tag{$\mathcal{P}_k$}$$

Show that :

1. For all $k \in \mathbb{N}$, $(\mathcal{P}_k)$ has at least one solution x_k.
2. The sequence $(x_k)_{n\in\mathbb{N}}$ is bounded.
3. Any cluster point of $(x_k)_{k\in\mathbb{N}}$ is a solution to $(\mathcal{P})$.
4. What can we say if F is strictly convex ?

Optimization – Exercises

Day 2

Exercise 1 (Convergence fixed step gradient descent algorithm).
For all $x \in \mathbb{R}^n$ we define the function f by

$$f(x) = \frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle,$$

where $A \in \mathcal{S}_n^{++}(\mathbb{R})$, with eigenvalues $(\lambda_i)_{1\leqslant i\leqslant n}$ verifying

$$0 < \lambda_1 \leqslant \ldots \leqslant \lambda_n,$$

and $b \in \mathbb{R}^n$. It has already been seen in exercise 4 that f admits a unique minimizer x^*, which is the solution to the linear system $Ax = b$.
The fixed step gradient descent algorithm is given by

$$\begin{cases} x_0 \in \mathbb{R}^n, \\ x_{k+1} = x_k - \gamma \nabla f(x_k). \end{cases}$$

Show the algorithm converges to x^* for any step $\gamma \in \left]0, \frac{2}{\lambda_n}\right[$. Give the step γ that ensures the fastest convergence.

Exercise 2 (Convergence of Uzawa method).
Let $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ be a differentiable α-strongly convex function and let $C \in \mathbb{R}^{m\times n}$, $d \in \mathbb{R}^m$. We propose to study the convergence of Uzawa method towards a solution to the following problem :

$$\begin{aligned} \underset{x\in\mathbb{R}^n}{\text{minimize}} \quad & f(x) \\ \text{subject to} \quad & Cx \leqslant d, \end{aligned} \qquad (\mathcal{P})$$

where the set $\{x \in \mathbb{R}^N \mid Cx \leqslant d\}$ is assumed to be nonempty. Let $\rho > 0$. Uzawa algorithm generates sequences $(x_k)_{k\in\mathbb{N}} \in (\mathbb{R}^n)^{\mathbb{N}}$ and $(\lambda_k)_{k\in\mathbb{N}} \in (\mathbb{R}^m)^{\mathbb{N}}$ according to the following iterations :

$$\boxed{\begin{cases} x_k = \underset{x\in\mathbb{R}^n}{\operatorname{argmin}}\ f(x) + \langle \lambda_k, Cx - d\rangle, \\ \lambda_{k+1} = \max\left(\lambda_k + \rho(Cx_k + d), 0\right). \end{cases}}$$

1. Explain why Problem $(\mathcal{P})$ admits a unique solution and why the algorithm is well defined.
2. (i) Write the Lagrangian $\mathcal{L}\colon \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}$ for Problem $(\mathcal{P})$.
   (ii) Show that for any $x \in \mathbb{R}^n$,

$$\left(\lambda^* = \underset{\lambda\in[0,+\infty)^m}{\operatorname{argmax}}\ \mathcal{L}(x, \lambda)\right) \iff \left((\forall \rho > 0) \quad \lambda^* = \mathrm{p}_+(\lambda + \rho(Cx - d))\right),$$

   where p_+ denotes the projection on $[0, +\infty)^m$.

(iii) Let (x^*, λ^*) be a saddle point of $\mathcal{L}$. Show that the following holds :

$$\begin{cases} \nabla f(x_k) - \nabla f(x^*) + C^\top(\lambda_k - \lambda) = 0 \\ \|\lambda_{k+1} - \lambda^*\| \leqslant \|\lambda_k - \lambda^* + \rho C(x_k - x^*)\|. \end{cases} \qquad (\star)$$

3. Using ($\star$), show the convergence of the sequence $(x_k)_{k\in\mathbb{N}}$ to x^* when ρ satisfies

$$0 < \rho < \frac{2\alpha}{\|C\|^2}. \qquad (\star\star)$$

Exercise 3 (Optimization with equality constraints).
Find the points (x, y, z) de $\mathbb{R}^3$ which belong to H_1 and H_2 and which are the closest to the origin.

$$\begin{aligned} (H_1) &: \ 3x + y + z = 5, \\ (H_2) &: \ x + y + z = 1. \end{aligned}$$

1. Write the problem as an optimization problem.
2. What can you say about existence of solutions ? Unicity ?
3. Solve the optimization problem using the Slater conditions.

Exercise 4 (Optimization with inequality constraints).
Solve the following optimization problem :

$$\begin{aligned} \underset{(x,y)\in\mathbb{R}^2}{\text{minimize}} \quad & x^4 + 3y^4 \\ \text{subject to} \quad & x^2 + y^2 \geqslant 1. \end{aligned}$$

Exercise 5 (Optimization with equality and inequality constraints).
Let $f : \mathbb{R}^k \longrightarrow \mathbb{R}$ be defined by

$$f(p_1, \dots, p_k) = \sum_{i=1}^{k} p_i^2.$$

Maximize f on the simplex Λ_k of $\mathbb{R}^k$

$$\Lambda_k := \left\{ p = (p_1, \dots, p_k) \in \mathbb{R}^k \mid p_i \geqslant 0 \text{ for all } i, \text{ and } \sum_{i=1}^{k} p_i = 1 \right\}.$$

Exercise 6 (Characterization of $\mathrm{SO}_n(\mathbb{R})$).
We denote $\mathrm{SO}_n(\mathbb{R}) = \{M \in \mathbb{R}^{n\times n} \mid M \text{ is orthogonal and } \det(M) = 1\}$ and $\mathrm{SL}_n(\mathbb{R}) = \{M \in \mathbb{R}^{n\times n} \mid \det(M) = 1\}$. Show $\mathrm{SO}_n(\mathbb{R})$ is exactly composed of the matrices of $\mathrm{SL}_n(\mathbb{R})$ which minimize the Euclidean norm of $\mathbb{R}^{n\times n}$, i.e.

$$\forall M \in \mathbb{R}^{n\times n}, \ \|M\| = \sqrt{\mathrm{Tr}(M^\top M)}.$$