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 :

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

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 $x \in \mathbb{R}^n$,

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

   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 :

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

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

$$0 < \rho < \frac{2\alpha}{\|C\|^2}. \tag{$\star\star$}$$

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

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

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

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

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.

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