Optimization for Machine Learning
Existence of minimizers / Convex sets

The exercises indicated with a star (⋆) are regarded as part of the syllabus. The stated results should be known, and it is highly recommended to practice by trying to solve them. **Last revised on October, 16th, 2019.**
The blue text denotes the addition of a question.
The red text denotes the correction of a typo.

Existence of minimizers

(⋆) **Exercise 1** (Limit inferior).

1. Let $(u_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ be a sequence, and assume that $\left(\liminf_{n\to+\infty} u_n\right) > -\infty$.

   Prove that $(u_n)_{n\in\mathbb{N}}$ is lower bounded, *i.e.* there is a constant $m \in \mathbb{R}$ such that $\forall n \in \mathbb{N}$, $u_n \geq m$.

2. Let $(u_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ be a sequence.
   a) Assume that $\lim_{n\to+\infty} u_n = \ell$ for some $\ell \in \mathbb{R}\cup\{\pm\infty\}$. Prove that $\liminf_{n\to+\infty} u_n = \ell$.
   b) Assume that there is a subsequence such that $\lim_{n\to+\infty} u_{\varphi(n)} = \ell$ for some $\ell \in \mathbb{R}\cup\{\pm\infty\}$. Prove that $\ell \geq \liminf_{n\to+\infty} u_n$.
   c) Prove that there exists a subsequence $(u_{\varphi(n)})_{n\in\mathbb{N}}$ such that $\lim_{n\to+\infty} u_{\varphi(n)} = \liminf_{n\to+\infty} u_n$.

3. Let $(u_n)_{n\in\mathbb{N}}$ and $(v_n)_{n\in\mathbb{N}}$ be two sequences in $\mathbb{R}\cup\{+\infty\}$. Prove that

$$\left(\liminf_{n\to+\infty} u_n\right) + \left(\liminf_{n\to+\infty} v_n\right) \leq \liminf\left(u_n + v_n\right).$$

   provided the left hand side is not $+\infty - \infty$. Give an example where the inequality is strict.

   Hint: *Consider for instance the sequence* $(-1)^n$, $n \in \mathbb{N}$.

(⋆) **Exercise 2** (Lower semi-continuous functions).
Let $f : \mathbb{R}^N \to \mathbb{R}\cup\{+\infty\}$.

1. Let $x \in \mathbb{R}^N$. Show that the following properties are equivalent[a)]:
   a) f is lower semi-continuous at x,
   b) **For all $t < f(x)$, there exists $r > 0$ such that for all $y \in B(x,r)$, $f(y) > t$.**

2. Show that the following properties are equivalent:
   a) f is lower semi-continuous on $\mathbb{R}^N$,
   b) For all $t \in \mathbb{R}$, the level set $\left\{ x \in \mathbb{R}^N \mid f(x) \leq t \right\}$ is closed.
   c) The epigraph $\left\{ (x,t) \in \mathbb{R}^N \times \mathbb{R} \mid f(x) \leq t \right\}$ is closed.

3. Prove that any continuous function $f : \mathbb{R}^N \to \mathbb{R}$ is lower semi-continuous.

(⋆) **Exercise 3** (Stability of l.s.c. functions).

1. Let $f_1, f_2 : \mathbb{R}^N \to \mathbb{R}\cup\{+\infty\}$ be two lower semi-continuous functions. Show that f_1+f_2 is lower semi-continuous.

2. Let $\{f_i\}_{i\in I}$ be a (finite or infinite) family of lower semi-continuous functions from $\mathbb{R}^N$ to $\mathbb{R}\cup\{+\infty\}$. Show that the function f defined by $f(x) \overset{\text{def}}{=} \sup_{i\in I}\left(f_i(x)\right)$ is lower semi-continuous.

3. Let $\{f_i\}_{1\leq i\leq k}$ be a **finite** family of lower semi-continuous functions from $\mathbb{R}^N$ to $\mathbb{R}\cup\{+\infty\}$. Show that the function f defined by $f(x) \overset{\text{def}}{=} \inf_{i\in I}\left(f_i(x)\right)$ is lower semi-continuous.

[a)] **If $f(x) < +\infty$ the second assertion may be reformulated as: For all $\varepsilon > 0$, there exists $r > 0$ such that for all $y \in B(x,r)$, $f(y) > f(x) - \varepsilon$.**

Exercise 4.
Do the following functionals have minimizers? Is it unique?

1. $f(x) = \sum_{i=1}^{N} x_i \log(x_i)$ on the domain $C = \left\{ x \in \mathbb{R}^N \mid x_i \geq 0, \sum_{i=1}^{N} x_i = 1 \right\}$.
2. $f(x) = \exp(x_1^2 - x_2^2)$ for $x = (x_1, x_2) \in \mathbb{R}^2$.
3. $f(x) = x_1^2 + x_2^2 + \exp(x_1^2 - x_2^2)$ for $x = (x_1, x_2) \in \mathbb{R}^2$.
4. $f(x) = \|Ax - y\|^2$ for $A \in \mathbb{R}^{m\times n}$, $y \in \mathbb{R}^m$.
5. $f(M) = \mathrm{Tr}(N^\top M)$ on $S_n^+(\mathbb{R}) \stackrel{\text{def}}{=} \{ M \in \mathbb{R}^{n\times n} \mid M^\top = M,\ M \succeq 0 \}$ for $N \in S_n^+(\mathbb{R})$?
6. $f(M) = \lambda_n(M)$ (maximal eigenvalue of M) on the domain $C = \left\{ M \in S_n^+(\mathbb{R}) \mid \max_{1\leq i,j\leq n} (|M_{i,j}|) = 1 \right\}$.

Convex sets

($\star$) **Exercise 5** (Minkowski sum).
Given two convex sets $C_1, C_2 \subset \mathbb{R}^N$, their sum (*or Minkowski sum*) is defined as

$$C_1 + C_2 \stackrel{\text{def}}{=} \{ x_1 + x_2 \mid x_1 \in C_1, x_2 \in C_2 \}.$$

1. What is $C_1 + C_2$ for $C_1 = \{a\}$ with $a \in \mathbb{R}^n$ and $C_2 \subset \mathbb{R}^n$ an arbitrary convex set?
   In the general case, observe that $C_1 + C_2 = \bigcup_{x_1 \in C_1} (\{x_1\} + C_2)$.
2. Draw the Minkowski sum $C_1 + C_2$ for the following sets
   a) $C_1 = [0, 1] \times \{0\}$, $C_2 = \{ (x, y) \in \mathbb{R}^2 \mid 0 \leq x \leq 1,\ |y| \leq x \}$.
   b) $C_1 = [0, 1]^2$, $C_2 = \bar{B}(0, r)$ (closed Euclidean ball with radius r) with $r < 1/2$.
3. Prove that $C_1 + C_2$ is convex.
4. Show that $C_1 + C_2$ is not necessarily closed, even if both C_1 and C_2 are closed.
   Hint: *Consider for instance* $C_1 = \{ (x, y) \in \mathbb{R}^2 \mid x \geq 0, y \geq 0, xy \geq 1 \}$ *and* $C_2 = \{0\} \times \mathbb{R}$.
5. Prove that $C_1 + C_2$ is closed if C_1 is compact and C_2 is closed.

Exercise 6 (Topological properties of convex sets).
Let $C \subset \mathbb{R}^N$ be a convex set.

1. Prove that $\overline{C}$, the closure of C, is convex (reminder: $\overline{C}$ is the set of all the limits of the converging sequences of elements of C).
2. Prove that $\mathrm{Int}(C)$, the interior of C, is convex (reminder: $\mathrm{Int}(C) \stackrel{\text{def}}{=} \{ x \in C \mid \exists r > 0, B(x, r) \subset C \}$).

Exercise 7 (Projection onto an affine subspace).

1. What is the characterization of the projection onto C when C is an affine (or linear) subspace of $\mathbb{R}^N$?
2. Let $C = \{ x \in \mathbb{R}^N \mid Ax = y \}$ where $A \in \mathbb{R}^{P\times N}$, $y \in \mathbb{R}^P$. We assume that $AA^\top$ is invertible. Compute the projection of $x_0 \in \mathbb{R}^N$ onto C.

($\star$) **Exercise 8** (Relative interior of a convex set).
The *relative interior* of a convex set is the interior of C when it is regarded as a subset of its affine hull $\mathrm{Aff}(C)$. In other words, $x \in \mathrm{ri}(C)$ if and only if

$$x \in \mathrm{Aff}(C) \quad \text{and} \quad \exists \mathbf{r} > \mathbf{0},\ \mathbf{B}(\mathbf{x}, \mathbf{r}) \cap \mathrm{Aff}(\mathbf{C}) \subset \mathbf{C}.$$

1. Find the relative interior of the following convex sets of $\mathbb{R}^3$:

$$C_1 = [0, 1]^3, \quad C_2 = \{ (x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 \leq 1, z = 1 \}, \quad C_3 = \mathbb{R} \times \{0\} \times \{0\}.$$

2. If the topological interior satisfies $\text{int}(C) \neq \emptyset$, what is the the relative interior $\text{ri}(C)$?

3. We want to prove that, if $C \neq \emptyset$, then $\text{ri}(C) \neq \emptyset$.
   a) Examine the case where C is a singleton, $C = \{x_0\}$.
   b) If C is not reduced to a singleton, show that there exists $m \geq 1$ and a family $\{c_0, \ldots, c_m\} \subset C$ such that $\text{Aff}(C) = \text{Aff}\{c_0, \ldots, c_m\}$ and $c_1 - c_0, \ldots, c_m - c_0$ are linearly independent.
   c) Prove that $\left(\frac{1}{m+1}\sum_{\mathbf{i=0}}^{m} c_i\right) \in \text{ri}(C)$.

4. If $x_1 \in \text{ri}(C)$ and $x_2 \in \overline{C}$, prove that $[x_1\,, x_2[\subset \text{ri}(C)$. ***Hint:*** *You may assume that* $\mathbb{R}^n = \text{Aff}(C)$ *and consider that* $\text{int}(C) \neq \emptyset$

Exercise 9 (Carathéodory's theorem).
Let $S \subset \mathbb{R}^N$, and let $\text{conv}(S)$ be its convex hull. Prove that any element of $\text{conv}(S)$ can be represented as a convex combination of (at most) $N+1$ elements of S.

Hint: *If x is a convex combination of $\sum_{i=1}^{k} \alpha_i s_i$ with $k > N+1$, then the family $\{s_i\}_{1\leq i\leq k}$ is affinely dependent, i.e. there exists coefficients $\beta_1, \ldots \beta_k$ not all zero such that $\sum_{i=1}^{k} \beta_i s_i = 0$ and $\sum_{i=1}^{k} \beta_i = 0$. Show that one can modify the α_i's so that one of them is zero.*

Exercise 10 (The Krein-Milman theorem (weak formulation)).
Let $C \subset \mathbb{R}^N$ be a convex set and $x_0 \in C$. We say that x_0 is *an extreme point of* C if there are no points $x_1, x_2 \in C$ such that $x_0 \in \,]x_1\,, x_2[$. We denote by $\text{extr}(C)$ the set of extreme points of C. We assume that C is *nonempty compact convex.*

1. Prove that C has at least one extreme point. ***Hint:*** *What can you say about the maximizers of the function $x \mapsto \|x\|^2$ on C?*

2. Let $y \in \mathbb{R}^N \setminus \{0\}$ and consider the set $F = \text{argmax}_{x\in C} \langle y, x\rangle$. Prove that F is nonempty compact convex, and that any extreme point of F is an extreme point of C.

3. Prove that $\overline{\text{conv}}(\text{extr}(C)) = C$, where $\overline{\text{conv}}$ denotes the closed convex hull.

   Hint: *What can you say if there is some $x \in C \setminus \overline{\text{conv}}(\text{extr}(C))$?*

   Note: *That result is known as the Krein-Milman theorem. In fact, in finite dimension (which is our case here), the closed convex hull $\overline{\text{conv}}$ can be replaced with the convex hull conv. In particular, any point of C is a convex combination of (at most) $N+1$ extreme points of C.*

4. Give a counter-example when C is not compact.

Exercise 11 (Projections onto convex sets).
Let $x \in \mathbb{R}^N$. Compute the projection of x onto

1. $C = \left\{ y \in \mathbb{R}^N \mid \|y\|_2 \leq 1 \right\}$ (unit ball).

2. $C = \left\{ y \in \mathbb{R}^N \mid \|y\|_\infty \leq 1 \right\}$ (unit cube).

3. $C = \left\{ (y,t) \in \mathbb{R}^{N-1} \times \mathbb{R} \mid \sum_{i=1}^{N-1} (y_i)^2 \leq 1 \text{ and } 0 \leq t \leq 1 \right\}$, for $N \geq 2$ (cylinder).