2 Convexity

Definition 2.1 A set $C \subset \mathcal{H}$ is **convex** if, for every $x, y \in C$

$$(\forall \alpha \in]0,1[) \quad \alpha x + (1-\alpha)y \in C. \tag{16}$$

A function f is **convex** if epi f is convex.

Proposition 2.2 $f\colon \mathcal{H} \to [-\infty, +\infty]$ *is convex if and only if*

$$(\forall x, y \in \operatorname{dom} f) \quad (\forall \alpha \in]0,1[) \qquad f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y). \tag{17}$$

Proof. First, we note that if f is identically $+\infty$, then dom $f = \varnothing$ if and only if epi $f = \varnothing$, so (17) is vacuously true. Now assume that dom $f \neq \varnothing$. Let (x, ξ) and (y, η) be in epi f and let $\alpha \in]0,1[$.
($\Rightarrow$) Assume that epi f is convex. Then

$$\alpha(x,\xi) + (1-\alpha)(y,\eta) = (\alpha x + (1-\alpha)y, \alpha\xi + (1-\alpha)\eta) \in \operatorname{epi} f. \tag{18}$$

Therefore, $f(\alpha x + (1-\alpha)y) \leq \alpha\xi + (1-\alpha\eta$. Taking the limit as $\xi \searrow f(x)$ and $\eta \searrow f(y)$ yields (17).
($\Leftarrow$) Assume that (17) holds. By definition, $f(x) \leq \xi$ and $f(y) \leq \eta$. So, by (17),

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y) \tag{19}$$

$$\leq \alpha\xi + (1-\alpha)\eta. \tag{20}$$

Therefore, $(\alpha x + (1-\alpha)y, \alpha\xi + (1-\alpha)\eta) \in \operatorname{epi} f$ which completes the proof. □

Definition 2.3 Let $\rho > 0$ and let $x \in \mathcal{H}$. A **closed ball** of radius ρ is $B(x;\rho) = \{z \in \mathcal{H} \mid \|x - z\| \leq \rho\}$.

Definition 2.4 Let $f\colon \mathcal{H} \to [-\infty, +\infty]$ and let $x \in \mathcal{H}$. x is a **local minimizer** of f if there exists $\rho > 0$ such that

$$(\forall z \in \mathcal{H} \cap B(x;\rho)) \quad f(x) \leq f(z). \tag{21}$$

x is a **global minimizer** of f if

$$(\forall z \in \mathcal{H}) \quad f(x) \leq f(z). \tag{22}$$

Fact 2.5 Let f be a convex and proper function. Then every local minimizer is a global minimizer.

Proof. This is left as an exercise (easier to prove after we learn about convex subdifferentials). □

Definition 2.6 Let $C \subset \mathcal{H}$ be nonempty.

(i) The **indicator function** of C is

$$\iota_C\colon \mathcal{H} \to [-\infty, +\infty] : x \mapsto \begin{cases} 0 & \text{if } x \in C \\ +\infty & \text{if } x \notin C. \end{cases} \tag{23}$$

(ii) Suppose that C is also closed. A **projection** of $x \in \mathcal{H}$ onto C is a solution to the minimization problem

$$\underset{z \in C}{\text{minimize}} \;\; \|x - z\|. \tag{24}$$

A solution to (24) is a "closest" point to x which resides in C.

Fact 2.7 Let $C \subset \mathcal{H}$ and let $x \in \mathcal{H}$.

(i) Without loss of generality, constrained optimization can be rephrased as unconstrained optimization via changing the objective function:

$$\inf_{x \in C} f(x) = \inf_{x \in \mathcal{H}} f(x) + \iota_C(x). \tag{25}$$

The objective function $f + \iota_C$ on the righthand side, although a bit fancier, allows us to rephrase the constraint on the lefthand side.

(ii) C is convex if and only if its indicator function ι_C is convex.

(iii) C is closed if and only if its indicator function ι_C is lsc.

(iv) Suppose that C is closed. Then a solution to (24) exists.

(v) Suppose that C is convex. If a solution to (24) exists, it is guaranteed to be unique.

The proofs of (ii) and (iii) follow from the fact that $\operatorname{epi} C = C \times [0, +\infty[$. Loosely speaking, the proof of (iv) follows from the Weierstraß theorem (compactness is achieved by intersecting C with $\{y \in \mathcal{H} \mid \|x - y\| \leq \eta\}$ for $\eta > 0$) and (v) follows from the fact that the norm is *strictly convex* – (a notion we have not yet defined, but we will see later in Definition 4.1).

Definition 2.8 Let $C \subset \mathcal{H}$ be nonempty, closed, and convex. In view of Fact 2.7(iv)–(v), for every $x \in \mathcal{H}$ there is a unique point, $\operatorname{Proj}_C(x) \in \mathcal{H}$, which solves (24). This implicitly defines the **projection operator** of C.

$$\operatorname{Proj}_C \colon \mathcal{H} \to \mathcal{H} \colon x \mapsto \operatorname{Proj}_C(x) \qquad \text{(solution to (24))} \tag{26}$$

Note: if $x \in C$, then $\operatorname{Proj}_C x = x$.

For all of the algorithms in this course, we will focus on functions from the following class

$$\Gamma_0(\mathcal{H}) = \big\{ f \colon \mathcal{H} \to \,]{-\infty}, +\infty] \;\big|\; f \;\text{ is proper, lower semicontinuous, and convex} \big\}. \tag{27}$$

The following functions live in $\Gamma_0(\mathcal{H})$:

(i) Exponentials: e^x

(ii) Log-barriers $f(x) = \begin{cases} -\ln(x) & \text{if } x > 0 \\ +\infty & \text{otherwise.} \end{cases}$

(iii) Any norm: $\|\cdot\|$ (e.g., $\|\cdot\|_1$ which promotes sparsity, $\|\cdot\|_{\text{nuclear}}$ which promotes low-rank)

(iv) Hinge-Loss, ReLU, KL-Divergence, ...

(v) Given a collection of functions $(f_i)_{i=1}^m$ in $\Gamma_0(\mathcal{H})$, we can remain in $\Gamma_0(\mathcal{H})$ via the following operations.

  (a) $\max\{f_1, \dots, f_m\}$

  (b) Positive linear combinations: $\lambda_1 f_1 + \cdots + \lambda_m f_m$, where $\{\lambda_i\}_{i=1}^m$ are positive.

  (c) Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be two finite-dimensional real vector spaces. Let $b \in \mathcal{H}_2$ and let $A\colon \mathcal{H}_1 \to \mathcal{H}_2$ be a linear operator (e.g., a matrix from $\mathbb{R}^n$ to $\mathbb{R}^m$). If $f_1 \in \Gamma_0(\mathcal{H}_2)$, then $g(x) = f_1(Ax + b) \in \Gamma_0(\mathcal{H}_1)$.

Exercise 2.9 The **Minkowski sum** of two subsets A, B of $\mathcal{H}$ is given by

$$A + B = \big\{ a + b \;\big|\; a \in A \;\text{ and }\; b \in B \big\}. \tag{28}$$

Assume that A and B are convex. Prove that $A + B$ is convex.

Exercise 2.10 Show that the norm $\|\cdot\|$ is convex using Definition 1.11.