

The following theorem is referenced a few times in the notes, so I will provide its statement here. Regretfully, this class does not have enough time to detail the topics of compact/closed/lsc. The following theorem is often used as a tool to ensure that a solution to an optimization problem exists.

Theorem 1.13 (Weierstraß) *Let $f: \mathcal{H} \rightarrow [-\infty, +\infty]$ be lower semicontinuous and let C be a compact subset of \mathcal{H} . Suppose that $C \cap \text{dom } f \neq \emptyset$. Then f achieves its infimum over C .*

2 Convexity

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

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

A function f is **convex** if $\text{epi } f$ is convex.

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

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

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

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

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),

$$f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y) \quad (19)$$

$$\leq \alpha \xi + (1 - \alpha)\eta. \quad (20)$$

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

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

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

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

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

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). \square

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

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

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

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

$$\underset{z \in C}{\text{minimize}} \quad \|x - z\|. \quad (24)$$

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:

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

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 $\text{epi } C = C \times [0, +\infty[$. 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 the interested student could research).

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, $\text{Proj}_C(x) \in \mathcal{H}$, which solves (24). This implicitly defines the **projection operator** of C .

$$\text{Proj}_C: \mathcal{H} \rightarrow \mathcal{H} : x \mapsto \text{Proj}_C(x) \quad (\text{solution to (24)}) \quad (26)$$

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

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

$$\Gamma_0(\mathcal{H}) = \{f: \mathcal{H} \rightarrow]-\infty, +\infty] \mid f \text{ is proper, lower semicontinuous, and convex}\}. \quad (27)$$

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 $(f_i)_{i \in I}$ 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: $\lambda_1 f_1 + \dots + \lambda_m f_m$, 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 $A: \mathcal{H}_1 \rightarrow \mathcal{H}_2$ 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

$$A + B = \{a + b \mid a \in A \text{ and } b \in B\}. \quad (28)$$

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 1.10.