Advanced Computational Statistics (Fall 2016)

J.P.Kim

Dept. of Statistics

Finally modified at October 12, 2016

Preface & Disclaimer

This note is a summary of the lecture Advanced Computational Statistics (M1399.000200) held at Seoul National University, Fall 2016. Lecturer was Jung-Ho Won, and the note was summarized by J.P.Kim, who is a Ph.D student. There are few textbooks and references in this course, which are following.

- *Convex Optimization Theory, Dimitri P. Bertsekas, 2009.*
- *Optimization, Kenneth Lange, 2013.*
- *Convex Optimization, S.Boyd & L.Vandenberghe, 2004.*

Also I referred to following books when I write this note. The list would be updated continuously.

- *Introduction to Mathematical Analysis (in Korean), Kim, Kim & Kye, 2012.*
- *Linear Algebra, S.H.Friedberg, 4th edition, 2003.*

If you want to correct typo or mistakes, please contact to: `joonpyokim@snu.ac.kr`

Chapter 1

Basic Concepts of Convex Analysis

1.1 Convex sets and functions

Definition 1.1.1. *A set $C \subseteq \mathbb{R}^n$ is **convex** if $\alpha x + (1-\alpha)y \in C$ for any $x, y \in C$ and $\alpha \in [0,1]$.*

Note that ϕ is convex by convention.

Proposition 1.1.2. *Let C and C_i be convex sets for $i \in I$. Then,*

(a) $\bigcap_{i \in I} C_i$ is also a convex set.

(b) $C_1 + C_2$ is a convex set.

(c) For any scalar λ, λC is a convex set. Also, for $\lambda_1, \lambda_2 > 0$, $(\lambda_1 + \lambda_2)C = \lambda_1 C + \lambda_2 C$ holds.

(d) $cl(C)$ and $int(C)$ are convex.

(e) For an affine function f, $f(C)$ or $f^{-1}(C)$ is convex.

Proof. (c) Convexity is trivial. Let $x \in (\lambda_1 + \lambda_2)C$. Then for some $y \in C$, $x = (\lambda_1 + \lambda_2)y$ holds. Since $\lambda_1 y \in \lambda_1 C$ and $\lambda_2 y \in \lambda_2 C$, we get $x \in \lambda_1 C + \lambda_2 C$. Thus we showed $(\lambda_1 + \lambda_2)C \subseteq \lambda_1 C + \lambda_2 C$. $\supseteq$ part is similar.

(d) Let $x, y \in cl(C)$. Then $\{x_k\}, \{y_k\} \subseteq C$ exist such that $x_k \to x$ and $y_k \to y$. Note that for any $\alpha \in [0,1]$ we get $\{\alpha x_k + (1-\alpha)y_k\} \subseteq C$, and so $\alpha x + (1-\alpha)y \in cl(C)$ from $\alpha x_k + (1-\alpha)y_k \to \alpha x + (1-\alpha)y$. Next, let $x, y \in int(C)$. Then there exists $r > 0$ such that $B(x,r) \subseteq C$ and $B(y,r) \subseteq C$. Note that $B(x,r) = \{x + z : \|z\| < r\}$. It's enough to show that $B(\alpha x + (1-\alpha)y, r) \subseteq C$. Now $B(\alpha x + (1-\alpha)y, r) = \{\alpha x + (1-\alpha)y + z : \|z\| < r\}$ and hence $\alpha x + (1-\alpha)y + z = \alpha \underbrace{(x+z)}_{\in B(x,r) \subseteq C} + (1-\alpha) \underbrace{(y+z)}_{\in B(y,r) \subseteq C} \in C$ for any z such that $\|z\| < r$.

(e) If $x, y \in f(C)$, $\exists x', y' \in C$ such that $x = f(x')$ and $y = f(y')$. Since f was affine, we get

$$\alpha x + (1-\alpha)y = \alpha f(x') + (1-\alpha)f(y') = f(\alpha x' + (1-\alpha)y') \in f(C)$$

from $\alpha x' + (1-\alpha)y' \in C$. Rest part is similar. □

Example 1.1.3 (Special convex sets)**.** In this example we see some examples of convex set.

(a) *Hyperplane* $\{x : a^T x = b\}$ is convex, for given a and b.

(b) *Half-space* $\{x : a^T x \le b\}$ is also convex.

(c) *Polyhedra*, $\{x : a_j^T x \le b_j,\ a_j \ne 0,\ b_j \in \mathbb{R},\ j = 1, 2, \cdots, r\}$ is intersection of half-spaces, and hence convex.

(d) C is *cone* if $\forall x \in C\ \lambda x \in C$ for any $\lambda > 0$. Note that, cone need not be convex, nor contain the origin. (See figure 1.1.) Rather, we consider *polyhedral cone* $\{x : a_j^T x \le 0,\ j = 1, 2, \cdots, r\}$, which contains the origin at the boundary. Polyhedral cone is convex.

(e) $S = \{x : a^T x = 0\}$ is a convex set, subspace of $\mathbb{R}^n$, a hyperplane, and a polyhedral cone.

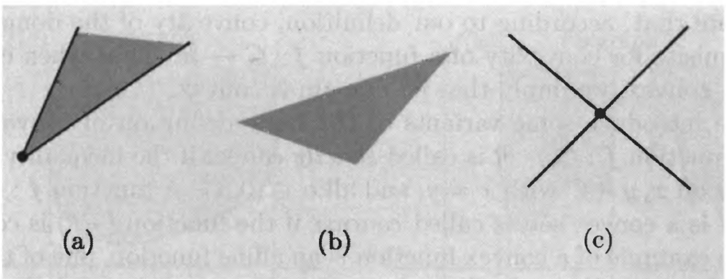


Figure 1.1: (a) Convex cone. (b) Convex cone which does not contain the origin. (c) Nonconvex cone, which consists of 2 lines.

Definition 1.1.4. *Let $C \subseteq \mathbb{R}^n$ be a convex set, and $f : C \to \mathbb{R}$ be a function. f is **convex** if*

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

*Remark: Domain is a convex set! Also, f is **strictly convex** if*

$$f(\alpha x + (1-\alpha)y) \le \alpha f(x) + (1-\alpha)f(y)\ \forall x, y \in C, x \ne y,\ \alpha \in (0, 1).$$

*Finally, f is **concave** if $-f$ is convex.*

From now on, without mention, C always denote a convex set in $\mathbb{R}^n$.

Example 1.1.5. (a) Affine function $f(x) = a^T x + b$ is both convex and concave.

(b) Any norm $f(x) = \|x\|$ is convex from *triangle inequality.*

Definition 1.1.6 (level set)**.** *Let $f : C \to \mathbb{R}$ be a convex function. Then for any given $\gamma \in \mathbb{R}$,*

(a) $\{x \in C : f(x) \le \gamma\}$ is called ***sublevel set*** *of f.*

(b) $\{x \in C : f(x) \ge \gamma\}$ is called ***superlevel set*** *of f.*

From now on, we will call a sublevel set as a ***level set*** *in short.*

Remark 1.1.7. It is known that if f is a convex function, then all of its level sets are convex. Note that converse does not hold: See figure 1.2.

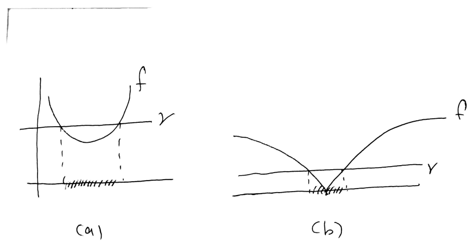


Figure 1.2: (a) Level set of convex function. (b) Even if all of level sets are convex, function need not be a convex one.

In many cases, it is convenient to allow function value be $\pm\infty$, or domain is $\mathbb{R}^n$. For this, we may consider a *extended real valued functions.* Then how to define convexity of such function $f : C \to [-\infty, \infty]$? The rest part of this section handles this issue.

Example 1.1.8. (Motivation for extension to $\bar{\mathbb{R}}$)

(a) We may deal with the function $f(x) = \sup_{i \in I} f_i(x)$. Its value may be ∞.

(b) "Conjugate function" will be handled in section 1.6. To define this notion, extension should be required. For example, conjugate function $f^*(y)$ of $f(x) = |x|$ is

$$f^*(y) = \begin{cases} 0 & |y| \le 1 \\ +\infty & o.w. \end{cases}.$$

(c) Consider $f(x) = 1/x$ on $(0, \infty)$. For optimization, closed domain is useful and convenient, so we may extend the domain to $[0, \infty]$. In here, $f(0) = \infty$ is reasonable extension.

Remark 1.1.9. Note that we can extend the domain of function $f : C \to \mathbb{R}$ to $\mathbb{R}^n$ as letting $f(x) = \infty$ if $x \notin C$. Thus allowing function to be extended real-valued, we can extend the domain of function. Then how to restrict the origin domain again? *Effective domain*, which is following, can be one answer.

Definition 1.1.10 (epigraph)**.** ***Epigraph*** *of function* $f : X \to \bar{\mathbb{R}}$ *is defined as*

$$epi(f) = \{(x, w) : x \in X,\ w \in \mathbb{R},\ f(x) \le w\}.$$

Note that w *is not allowed to be* $\pm\infty$*.*

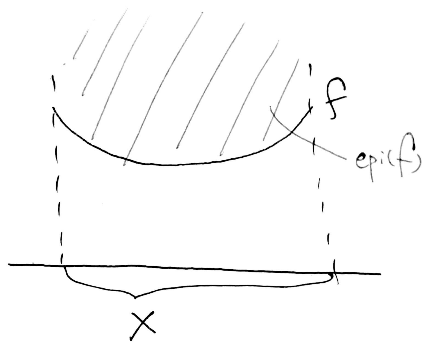


Figure 1.3: Epigraph of a function

Proposition 1.1.11. *Epigraph of convex function is a convex set.*

Proof. Easy. □

Definition 1.1.12 (effective domain)**.** *Let* $f : X \to \bar{\mathbb{R}}$ *be a function.* ***Effective domain*** *of* f *is defined as*

$$dom(f) = \{x \in X : f(x) < \infty\}.$$

There are some remarks.

Remark 1.1.13.

(a) Since we usually deal with a convex function f, the point whose functional value is $-\infty$ is out of interest.

(b) Note that

$$dom(f) = \{x \in \mathbb{R}^n : \exists w \in \mathbb{R}\ s.t.\ (x, w) \in epi(f)\},$$

so it is "projection of $epi(f)$ onto $\mathbb{R}^n$. If we want to handle real valued function, we can think restriction on $dom(f)$. Or, as mentioned above, we can enlarge domain from X to $\mathbb{R}^n$. Extended or restricted functions have the same epigraph.

Example 1.1.14.

(a) Consider a function $f : [0, \infty) \to [-\infty, \infty]$ defined as

$$f(x) = \begin{cases} \dfrac{1}{x} & x \neq 0 \\ +\infty & x = 0 \end{cases}.$$

Then $dom(f) = (0, \infty)$, and

$$epi(f) = \{(x, y) : 0 < x < \infty, y \geq 1/x.\}$$

(b) Suppose that $f(x) = -\infty$ for some $x \in X$. Then its epigraph $epi(f)$ may contain a vertical line.

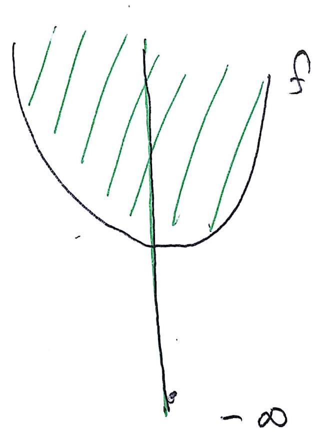


Figure 1.4: Epigraph of nonproper function

Definition 1.1.15 (proper function)**.** *Let $f : X \to \bar{\mathbb{R}}$. f is a* ***proper function*** *if*

(1) $f(x) < \infty$ for at least one $x \in X$, and

(2) $f(x) > -\infty$ for all $x \in X$.

Remark 1.1.16. Note that, f is proper function $\Leftrightarrow$ $epi(f)$ is nonempty and does not contain vertical line.

Now we can extend the definition of convex function *to extended real valued function.*

Definition 1.1.17. *$f : C \to \bar{\mathbb{R}}$ is a convex function if epi(f) is a convex subset of* $\mathbb{R}^{n+1}$.

Remark 1.1.18. Note that this definition satisfies followings.

(1) $dom(f)$ is convex.

(2) All of level sets are convex.

(3) If $f(x) < \infty\ \forall x$ or $f(x) > -\infty\ \forall x$, it satisfies Jensen's inequality.

Definition 1.1.19 (Indicator function)**.** *Let $X \subseteq \mathbb{R}^n$ be a set. An **indicator function** δ_X of X is defined as*

$$\delta_X(x) = \begin{cases} 0 & x \in X \\ +\infty & o.w. \end{cases}.$$

Note that effective domain of δ_X is X. Also, note that

$$X \text{ is (strictly) convex set } \Leftrightarrow \delta_X \text{ is (strictly) convex function.}$$

Also, if $X \neq \phi$, δ_X is proper.

Remark 1.1.20. Now we can give a correspondence between convex sets and convex functions. Epigrpah of convex function is convex set, and indicator of convex set is convex function.

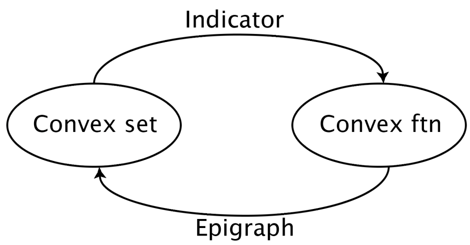


Figure 1.5: Correspondence between convex sets and functions

Now, we are ready for extension of convexity to nonconvex domain.

Definition 1.1.21. *Let C be a convex set, and $C \subseteq X \subseteq \mathbb{R}^n$. Then $f : X \to \bar{\mathbb{R}}$ is **convex over C** if $f|_C : C \to \bar{\mathbb{R}}$ (restriction on C of f) is convex function.*

1.1.1 Closedness and Semicontinuity

Definition 1.1.22. *A function $f : X \to \bar{\mathbb{R}}$ is a closed function if its epigraph epi(f) is closed set.*

It is reasonable definition because of correspondence between sets and functions. In Appendix A, we defined lower and upper semicontinuity of function. There is an important relationship between two notions of function. In fact, *they are equivalent.*

Theorem 1.1.23. *Let $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ be a function. Then TFAE*[1]

(i) For any $\gamma \in \mathbb{R}$, $V_\gamma = \{x : f(x) \le \gamma\}$ is closed.

(ii) f is lower semicontinuous function.

(iii) f is closed function. (epi(f) is closed)

Proof. If $f(x) \equiv \infty$, it is trivial, so assume not.

(i)⇒(ii): Suppose that $\bar{x}$ and a sequence $\{x_k\}$ exist such that $x_k \xrightarrow[k\to\infty]{} \bar{x}$ and $f(x) > \liminf_{k\to\infty} f(x_k)$. Then $\exists\gamma$ such that $f(\bar{x}) > \gamma > \liminf_{k\to\infty} f(x_k)$. Then there is a subsequence $\{x_{kj}\}$ which satisfies

$$f(x_{kj}) \le \gamma \ \forall j,$$

by definition of lim inf. Hence $V_\gamma := \{x : f(x) \le \gamma\} \supseteq \{x_{kj}\}$, and from closedness of V_γ, $\bar{x} \in V_\gamma$ should be held, which yields contradiction.

(ii) ⇒ (iii): Choose a sequence $\{(x_k, w_k)\} \subseteq epi(f)$ such that $(x_k, w_k) \to (\bar{x}, \bar{w})$. Then since f is l.s.c.,

$$f(\bar{x}) \le \liminf_{k\to\infty} f(x_k) \le \liminf_{k\to\infty} w_k$$

holds by definition of epigraph. Thus we get

$$f(\bar{x}) \le \bar{w}$$

by letting $k \to \infty$. Therefore $(\bar{x}, \bar{w}) \in epi(f)$ holds.

(iii) ⇒ (i): Note that $(x, \gamma) \in epi(f) \Leftrightarrow x \in V_\gamma$. Let $\gamma \in \mathbb{R}$, and $\{x_k\} \subseteq V_\gamma$ be a sequence converging to $\bar{x}$. Then $(x_k, \gamma) \in epi(f)$ and $(x_k, \gamma) \xrightarrow[k\to\infty]{} (\bar{x}, \gamma)$ hold, which imply $(\bar{x}, \gamma) \in epi(f)$ since $epi(f)$ is closed. Therefore $\bar{x} \in V_\gamma$ is obtained. □

Remark 1.1.24. We will often use the condition that a function is *closed*, rather than *lower semicontinuity*, even though they are equivalent on $\mathbb{R}^n$. It's because closedness of epigraph is more convenient to handle, due to the 'domain dependency' of semicontinuity. For example,

[1] The followings are equivalent.

consider a function

$$f : \mathbb{R} \to (-\infty, \infty],\ f(x) = \begin{cases} 0 & 0 < x < 1 \\ \infty & o.w. \end{cases}.$$

Then its epigraph is $epi(f) = (0,1) \times [0, \infty)$ so it is not closed, nor lower semicontinuous. However, if we restrict the domain,

$$\tilde{f} : (0,1) \to (-\infty, \infty],\ \tilde{f}(x) = 0$$

is lower semicontinuous, while its epigraph does not change, which means that $\tilde{f}$ is not closed. For this reason, we often consider the epigraph while we deal with closedness or semicontinuity of function.

Then our question is: Cannot we think similar thing as theorem 1.1.23 for a function on restricted domain? Next theorem gives the answer.

Proposition 1.1.25. *Let $f : X \to \bar{\mathbb{R}}$ and suppose that $dom(f)$ is closed, and f is l.s.c. at x for any $x \in dom(f)$. Then, f is closed.*

Proof. Similar as 1.1.23. □

Example 1.1.26. Let $X \subseteq \mathbb{R}^n$. Then,

(a) Indicator δ_X of X is closed iff X is closed.

(b) Let

$$f_X(x) = \begin{cases} f(x) & x \in X \\ \infty & o.w. \end{cases} \qquad \text{(“extension to the whole domain”)}$$

Then f_X is closed iff X is closed.

Proof. (From HW1) Let δ_X be indicator of X. Then

$$epi(\delta_X) = \{(x, w) : x \in X, w \geq 0\} = X \times [0, \infty)$$

is closed iff X is closed. Next,

$$epi(f_X) = \{(x, w) : x \in X, f(x) \leq w\}.$$

If $epi(f_X)$ is closed, $\forall (x_k, w_k) \to (\bar{x}, \bar{w})$ s.t. $\{(x_k, w_k)\} \subseteq epi(f_X)$, $(\bar{x}, \bar{w}) \in epi(f_X)$. Thus $\bar{x} \in X$. Note that $\forall x_k \to \bar{x}\ \exists w_k$ s.t. $(x_k, w_k) \to (\bar{x}, \bar{w})$ and $\{(x_k, w_k)\} \subseteq epi(f_X)$. Conversely, if X is closed, $\forall(x_k, w_k) \to (\bar{x}, \bar{w})$ since $x_k \to \bar{x}$ so $\bar{x} \in X$ and $f(x_k) \le w_k \Leftrightarrow f(\bar{x}) \le \bar{w}$ so $(\bar{x}, \bar{w}) \in epi(f_X)$. (continuity of f is used) Hence $epi(f_X)$ is closed. □

In optimization, we usually consider a *proper, convex and closed* functions. Following proposition says that 'proper' condition is needed to make the function reasonable.

Proposition 1.1.27. *Improper closed convex function* ***cannot*** *take a finite value anywhere.*

Proof. Let $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ be an improper closed convex function. Suppose that $\exists x$ such that $f(x) \in \mathbb{R}$ (i.e., f has a finite value). Then $f \not\equiv \infty$ and so $\exists \bar{x}$ s.t. $f(\bar{x}) = -\infty$. Define a sequence $\{x_k\}$ as

$$x_k = \frac{k-1}{k}x + \frac{1}{k}\bar{x}.$$

Note that $x_k \to x$ as $k \to \infty$. By convexity,

$$f(x_k) \le \frac{k-1}{k}f(x) + \frac{1}{k}f(\bar{x}) = -\infty,$$

so we get $\forall k$ $f(x_k) = -\infty$. Now by closedness, f is lower semicontinuous, and so

$$f(x) \le \liminf_{k\to\infty} f(x_k) = -\infty,$$

which yields contradiction. □

Remark 1.1.28. Note that by previous proposition, improper closed convex can have only the form as

$$f(x) = \begin{cases} -\infty & x \in dom(f) \\ \infty & o.w. \end{cases}.$$

1.1.2 Operations that preserve convexity of functions

Following operations preserve convexity.

(a) Composition with a linear transform, $f(Ax)$, where f: convex and A is $m \times n$ matrix. (It also preserves closedness)

(b) Summation or positive scalar multiplication, $\lambda_1 f_1(x) + \cdots + \lambda_m f_m(x)$ where f_i's are convex and $\lambda_i > 0$.

(c) Taking sup (See proposition 1.1.29)

(d) Taking partial minimum. If $f(x, z)$ is convex in (x, z), then $x \mapsto \inf_z f(x, z)$ is convex (Will be shown at section 3.3.).

Proposition 1.1.29. *Let $f_i : \mathbb{R}^n \to \mathbb{R}$ be convex functions where $i \in I$. Then*

$$f(x) := \sup_{i \in I} f_i(x)$$

is also convex.

Proof. We use the definition of convexity of *extended real-valued* function. Note that

$$(x, w) \in epi(f) \Leftrightarrow f(x) \le w \Leftrightarrow f_i(x) \le w \ \forall i \in I \Leftrightarrow (x, w) \in epi(f_i) \ \forall i \in I \Leftrightarrow (w, x) \in \bigcap_{i \in I} epi(f_i)$$

so we obtain

$$epi(f) = \bigcap_{i \in I} epi(f_i),$$

which yields the desired result. □

Remark 1.1.30. Note that $epi(f) = \bigcap_{i \in I} epi(f_i)$ also implies that f is closed, i.e., taking supremum preserves closedness as well as convexity.

1.1.3 Differentiable convex functions

In this subsection we deal with *differentiable* convex functions. Since we can define a gradient of function, There are some more things that we can say.

Proposition 1.1.31. *Let $C \subseteq \mathbb{R}^n$ be a nonempty convex set, and $f : \mathbb{R}^n \to \mathbb{R}$ be differentiable over an open set containing C. Then*

(a) f is convex over $C \Leftrightarrow f(z) \ge f(x) + \langle \nabla f(x), z - x \rangle \ \forall x, z \in C$

(b) f is (strictly) convex over $C \Leftrightarrow f(z) \ge f(x) + \langle \nabla f(x), z - x \rangle \ \forall x, z \in C$ s.t. $x \ne z$

Proof. Only a proof for (a) would be given.

$\Leftarrow$) Let $x, y \in C$, $\alpha \in [0, 1]$, and $z = \alpha x + (1 - \alpha)y$. Then by the assumption,

$$f(x) \ge f(z) + \langle \nabla f(z), x - z \rangle$$

$$f(y) \ge f(z) + \langle \nabla f(z), y - z \rangle$$

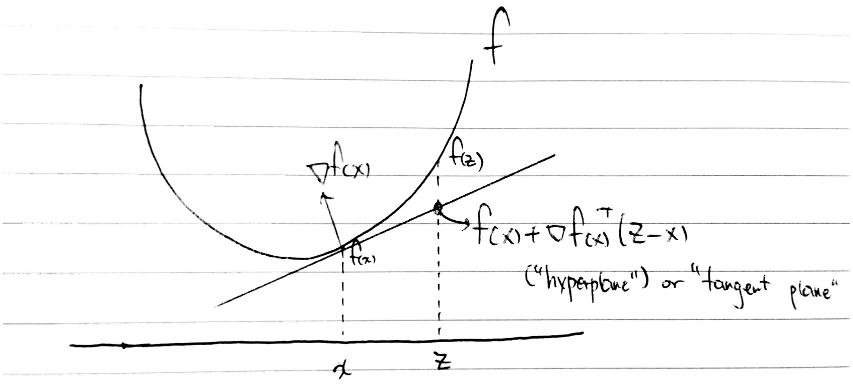


Figure 1.6: Convex differentiable function. See proposition 1.1.31.

holds. Thus, we get

$$\alpha f(x)+(1-\alpha y) \geq f(z)+\langle \nabla f(z), \alpha(x-z)+(1-\alpha)(y-z)\rangle = f(z)+\langle \nabla f(z), \underbrace{\alpha x+(1-\alpha)y}_{=z} - z\rangle$$

and therefore

$$\alpha f(x)+(1-\alpha)y \geq f(z) = f(\alpha x+(1-\alpha)y).$$

$\Rightarrow$) Let $x, y \in C$ and $x \neq y$. Define

$$g(\alpha)=\frac{f(x+\alpha(z-x))-f(x)}{\alpha} \text{ for } \alpha \in (0,1]. \qquad (\text{"Average rate on the direction of } z-x\text{"})$$

Then we get

$$\lim_{\alpha \searrow 0} g(\alpha) = \langle \nabla f(x), z-x\rangle \qquad (\text{"Directional derivative"})$$

and

$$g(1)=f(z)-f(x).$$

Thus if we can show that g is monotonely increasing,

$$g(1) \geq \lim_{\alpha \searrow 0} g(\alpha)$$

holds, which is the desired result. So our claim is:

Claim. g is monotonely increasing.

Choose $0<\alpha_1<\alpha_2<1$. Then

$$f(x+\alpha_1(z-x)) = f\left(\frac{\alpha_1}{\alpha_2}(x+\alpha_2(z-x))+\left(1-\frac{\alpha_1}{\alpha_2}\right)x\right)$$

$$\leq \frac{\alpha_1}{\alpha_2} f(x+\alpha_2(z-x)) + \left(1 - \frac{\alpha_1}{\alpha_2}\right) f(x)$$

so

$$\frac{f(x+\alpha_1(z-x)) - f(x)}{\alpha_1} \leq \frac{f(x+\alpha_2(z-x)) - f(x)}{\alpha_2}$$

is obtained. □

Remark 1.1.32. Proposition 1.1.31 has some siginficant consequences.

(1) If $f : \mathbb{R}^n \to \mathbb{R}$ is a differentiable convex function, then for x^* s.t. $\nabla f(x^*) = 0$ ("critical point") we get

$$f(x) \geq f(x^*) + \langle \nabla f(x^*), x - x^* \rangle \ \forall x$$

and hence

$$x^* \in \arg\min_{x \in \mathbb{R}^n} f(x). \qquad \text{("Unconstrained Optimization")}$$

(2) If $\langle \nabla f(x^{**}), z - x^{**} \rangle \geq 0 \ \forall z \in C$ holds, then we get

$$f(z) \geq f(x^{**}) + \langle \nabla f(x^{**}), z - x^{**} \rangle \geq f(x^{**}) \ \forall z \in C$$

so

$$x^{**} \in \arg\min_{x \in C} f(x). \qquad \text{("Constrained Optimization")}$$

(3) In fact, converse of (2) also holds. In other words, if $x^{**} \in C$ minimizes f over C, then $\langle \nabla f(x^{**}), z - x^{**} \rangle \geq 0 \ \forall z \in C$. To see this, assume that $\langle \nabla f(x^{**}), z - x^{**} \rangle < 0$ for some $z \in C$. Then since $\langle \nabla f(x^{**}), z - x^{**} \rangle$ is a directional derivative, we get

$$\lim_{\alpha \searrow 0} \frac{f(x^{**} + \alpha(z - x^{**})) - f(x^{**})}{\alpha} = \langle \nabla f(x^{**}), z - x^{**} \rangle < 0,$$

so for small α, we get

$$f(x^{**} + \alpha(z - x^{**})) < f(x^{**}),$$

which yields contradiction to minimization assumption of x^{**}.

(4) Later, proposition 1.1.31 will be extended to *subdifferential functions* using *subgradients.*

Proposition 1.1.33 (Projection Theorem)**.** *Let $C \subseteq \mathbb{R}^n$ be a nonempty closed convex set, and*

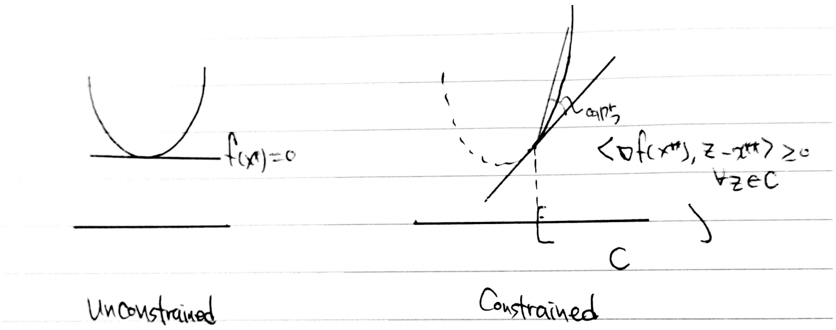


Figure 1.7: Unconstrained and Constrained Optimization

$z \in \mathbb{R}^n$ be a vector. Then there is a unique vector x^* such that

$$\|z - x^*\| \leq \|z - x\| \ \forall x \in C.$$

In this case, we denote $x^* = \mathcal{P}_C(z) = \underset{x \in C}{\arg\min} \|z - x\|$. Furthermore,

$$x^* = \mathcal{P}_C(z) \ \Leftrightarrow \ \langle z - x^*, x - x^* \rangle \leq 0 \ \forall x \in C.$$

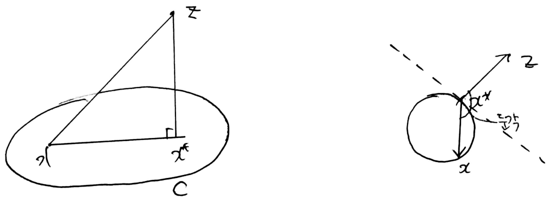


Figure 1.8: Projection Theorem

Proof. (Existence) Let $\tilde{C} = C \cap \{x : \|z - x\| \leq \|z - w\|\}$ for some $w \in C$. Since we think minimization, we get

$$\min_{x \in C} f(x) = \min_{x \in \tilde{C}} f(x) \ . \qquad \text{("Restriction to bounded ball")}$$

(See figure 1.9) Then since $\tilde{C}$ is compact, by max-min theorem, we get $\exists x^* = \underset{x \in \tilde{C}}{\arg\min} f(x)$.

(Uniqueness) Let x_1^*, x_2^* be minimizers. Then by the fact that will be shown, we get

$$\langle z - x_1^*, x_2^* - x_1^* \rangle \leq 0$$

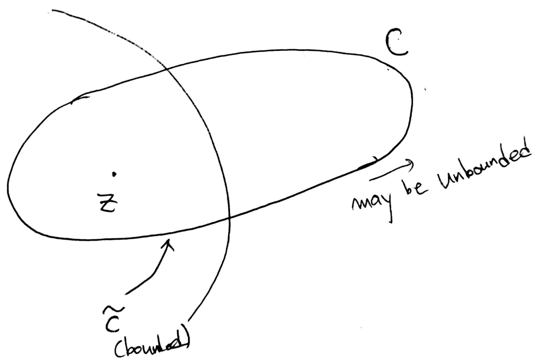


Figure 1.9: Restriction to bounded ball

$$\langle z - x_2^*, x_1^* - x_2^* \rangle \le 0,$$

which implies $\langle x_2^* - x_1^*, x_2^* - x_1^* \rangle \le 0$, and hence $x_1^* = x_2^*$.

(Rest part) Let $f(x) = \|z - x\|^2/2$. Then from previous theorem,

$$x^* = \arg\min_{x \in C} f(x) \;\Leftrightarrow\; \langle \nabla f(x^*), x - x^* \rangle \ge 0 \;\forall x \in C$$

holds, so from $\nabla f(x^*) = x^* - z$, we get $\langle z - x^*, x - x^* \rangle \le 0$. □

Now consider a $\mathcal{C}^2$ function.

Proposition 1.1.34. *Let $C \subseteq \mathbb{R}^n$ be a nonempty convex set, and $f : \mathbb{R}^n \to \mathbb{R}$ be twice continuously differentiable function over an open set that contains C. Then,*

(a) $\nabla^2 f(x) \ge 0 \;\forall x \in C \;\Rightarrow f$: convex over C.

(b) $\nabla^2 f(x) > 0 \;\forall x \in C \;\Rightarrow f$: strictly convex over C.

(c) If C is open and f is convex over C, then $\nabla^2 f(x) \ge 0 \;\forall x \in C$.

Proof. By Taylor's theorem, for any $x, y \in C$, there is $\alpha \in [0, 1]$ such that

$$f(y) = f(x) + \nabla f(x)^T (y - x) + \frac{1}{2}(y - x)\nabla^2 f\left(x + \alpha(y - x)\right)(y - x).$$

Thus by theorem 1.1.31 we get (a) and (b). For (c), assume not. Then $\exists x \in C$ and $\exists z \in \mathbb{R}^n$ such that $z^T \nabla^2 f(x) z < 0$. WLOG z has very small norm. Then $x + z \in C$ from openness of C, and since $\nabla^2 f$ is continuous, for any $\alpha \in [0, 1]$, we get $z^T f(x + \alpha z) z < 0$. Using Taylor theorem again, we can yield contradiction. □

Remark 1.1.35. In (c), if open condition of C is omitted, then the assertion does not hold. For example, let $C = \{(x_1, 0) : x_1 \in \mathbb{R}\}$ and $f(x_1, x_2) = x_1^2 - x_2^2$. Then f is convex over C but

$$\nabla^2 f(x_1, x_2) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$$

is not positive definite.

1.2 Convex and Affine hulls

In this section, our goal is "convexification" of nonconvex sets. Let $x_1, \cdots, x_k \in \mathbb{R}^n$, $\alpha_1, \cdots, \alpha_k \in \mathbb{R}$ and $S \subseteq \mathbb{R}^n$ be a nonempty set. We can summarize definitions and facts about hulls as table 1.1.

	Linear	Affine	Convex
combination	$\sum_{i=1}^k \alpha_i x_i$	$\sum_{i=1}^k \alpha_i x_i$, where $\sum_{i=1}^k \alpha_i = 1$	$\sum_{i=1}^k \alpha_i x_i$, where $\sum_{i=1}^k \alpha_i = 1$, $\forall \alpha_i \geq 0$.
set	X is a linear set $\Leftrightarrow X = \left\{\sum_{i=1}^k \alpha_i x_i : x_i \in X\right\}$ $\Leftrightarrow X$ is a subspace	X is an affine set $\overset{(a)}{\Leftrightarrow} X = \left\{\sum_{i=1}^k \alpha_i x_i : x_i \in X, \sum_{i=1}^k \alpha_i = 1\right\}$	X is a conex set $\overset{(e)}{\Leftrightarrow} X = \left\{\sum_{i=1}^k \alpha_i x_i : x_i \in X, \sum_{i=1}^k \alpha_i = 1, \forall \alpha_i \geq 0\right\}$
hull generation	$\mathrm{lin}(S) = \bigcap(\text{linear} \supseteq S) = \left\{\sum_{i=1}^k \alpha_i x_i : x_i \in S\right\}$	$\mathrm{aff}(S) = \bigcap(\text{affine} \supseteq S) \overset{(b)}{=} \left\{\sum_{i=1}^k \alpha_i x_i : x_i \in S, \sum_{i=1}^k \alpha_i = 1\right\}$	$\mathrm{conv}(S) = \bigcap(\text{convex} \supseteq S) \overset{(f)}{=} \left\{\sum_{i=1}^k \alpha_i x_i : x_i \in S, \sum_{i=1}^k \alpha_i = 1, \forall \alpha_i \geq 0\right\}$
independence	$x_1, \cdots, x_k$ is linearly indep if $\sum_{i=1}^k \alpha_i x_i = 0 \Rightarrow \forall \alpha_i = 0$	$x_1, \cdots, x_k$ is affinely indep if $x_1 + \langle x_2 - x_1, \cdots, x_k - x_1\rangle$ has dimension $k-1$ ((c))	N/A
cardinality	$\forall x \in \mathrm{lin}(S)$ can be represented as a linear combination of no more than n points from S $(k \leq n)$	(d) $\forall x \in \mathrm{aff}(S)$ can be represented as a affine combination of no more than $n+1$ points from S $(k \leq n+1)$	(g) $\forall x \in \mathrm{conv}(S)$ can be represented as a convex combination of no more than $n+1$ points from S $(k \leq n+1)$

Table 1.1: Linear, affine, and convex hull

Remark 1.2.1. Some remarks or proofs about table 1.1.

(a) Note that for some $x_0 \in X$, we get $X = x_0 + (X - x_0)$, and $X - x_0$ is a subspace. Then we get

$$X = x_0 + \left\{ \sum_{i=1}^{k} \alpha_i (x_i - x_0) : x_i \in X,\ i = 1, 2, \cdots, k \right\} = \left\{ \sum_{i=0}^{k} \alpha_i x_i : x_i \in X,\ i = 1, 2, \cdots, k \right\}$$

letting $\alpha_0 = 1 - \alpha_1 - \cdots - \alpha_k$.

(b) Take $x_0 \in S$ and A be an affine set that contains S. Then $A = x_0 + (A - x_0)$ holds, and $A - x_0$ is a subspace that contains $S - x_0$. Thus $\text{aff}(S) = x_0 + \text{span}(S - x_0)$. Now from

$$\text{span}(S - x_0) = \left\{ \sum_{i=1}^{k} \alpha_i (x_i - x_0) : x_i \in S,\ i = 1, 2, \cdots, k \right\},$$

we get

$$\text{aff}(S) = \left\{ \sum_{i=0}^{k} \alpha_i x_i : x_i \in S,\ i = 1, 2, \cdots, k \right\}.$$

(c) By index changing, we get

$$x_1 + \text{span}(x_2 - x_1, \cdots, x_k - x_1) = x_k + \text{span}(x_1 - x_k, \cdots, x_{k-1} - x_k) = x_k + \text{span}(x_1 - x_k, \cdots, x_k - x_k).$$

(In the table, $\text{span}(\cdot, \cdot)$ is represented as $\langle \cdot, \cdot \rangle$) Now by definition, $x_1, \cdots, x_k$ are affinely independent if $x_i - x_k,\ i = 1, 2, \cdots, k-1$ are linearly independent, i.e.,

$$\sum_{i=1}^{k-1} \alpha_i (x_i - x_k) = 0 \ \Rightarrow \alpha_1 = \alpha_2 = \cdots = \alpha_{k-1} = 0.$$

Now letting $\alpha_k = -(\alpha_1 + \cdots + \alpha_{k-1})$, we get

$$\sum_{i=1}^{k} \alpha_i = 0 \text{ and } \sum_{i=1}^{k} \alpha_i x_i = \sum_{i=1}^{k-1} \alpha_i (x_i - x_k).$$

Therefore, affinely independent condition is equivalent to

$$\sum_{i=1}^{k} \alpha_i x_i = 0,\ \sum_{i=1}^{k} \alpha_i = 0 \Rightarrow \alpha_1 = \cdots = \alpha_k = 0$$

holds, which means that

$$\sum_{i=1}^{k}\alpha_i\begin{pmatrix}x_i\\1\end{pmatrix}=0$$

has a unique solution. In other words, $x_1,\cdots,x_k$ is affinely independent iff k vectors $\begin{pmatrix}x_1\\1\end{pmatrix},\cdots,\begin{pmatrix}x_k\\1\end{pmatrix}$ in $\mathbb{R}^{n+1}$ are linearly independent.

(d) Since affinely independent condition in $\mathbb{R}^n$ is equivalent to linearly independent condition in $\mathbb{R}^{n+1}$, at most $n+1$ points determine $\mathrm{aff}(S)$.

(e) $\Leftarrow$ part is trivial. For $\Rightarrow$ part, let $x_1,\cdots,x_k\in X$ and $\alpha_1,\cdots,\alpha_k\in\mathbb{R}$, $\sum\alpha_i=1$, and $\alpha_i\ge 0$. As summation is 1, at least one α_i is positive. WLOG $\alpha_1>0$. Now

$$\begin{aligned}
y_2&=\frac{\alpha_1}{\alpha_1+\alpha_2}x_1+\frac{\alpha_2}{\alpha_1+\alpha_2}x_2\in X\ (\because \text{convexity})\\
y_3&=\frac{\alpha_1+\alpha_2}{\alpha_1+\alpha_2+\alpha_3}y_2+\frac{\alpha_3}{\alpha_1+\alpha_2+\alpha_3}x_3=\frac{\alpha_1x_1+\alpha_2x_2+\alpha_3x_3}{\alpha_1+\alpha_2+\alpha_3}\in X\\
&\vdots\\
y_k&=\frac{\alpha_1+\cdots+\alpha_{k-1}}{\alpha_1+\cdots+\alpha_k}y_{k-1}+\frac{\alpha_k}{\alpha_1+\cdots+\alpha_k}x_k=\frac{\alpha_1x_1+\cdots+\alpha_kx_k}{\alpha_1+\cdots+\alpha_k}\in X
\end{aligned}$$

hold, and from $\sum\alpha_i=1$, we get $y_k=\alpha_1x_1+\cdots+\alpha_kx_k\in X$.

(f) Let

$$T:=\left\{\sum_{i=1}^{k}\alpha_ix_i : x_i\in S,\ \sum_{i=1}^{k}\alpha_i=1,\ \forall\alpha_i\ge 0\text{ for some }k\right\}.$$

Then clearly T is convex, so $T\supseteq\mathrm{conv}(S)$. Now if $S'\supseteq S$ is a convex set, then by (e)

$$S'=\left\{\sum_{i=1}^{k}\alpha_ix_i : x_i\in S',\ \sum_{i=1}^{k}\alpha_i=1,\ \forall\alpha_i\ge 0\text{ for some }k\right\}$$

and it contains T from $S'\supseteq S$. Take intersection on S' and we obtain $\mathrm{conv}(S)\supseteq T$. Therefore $\mathrm{conv}(S)=T$.

(g) It is the result of *Carathéodory theorem*; any $x\in\mathrm{conv}(S)$ can be represented as a convex combination of at most $n+1$ points from S. It means that for

$$C_k:=\left\{\sum_{i=1}^{k}\alpha_ix_i : x_i\in S,\ \alpha_i\ge 0,\ \sum_{i=1}^{k}\alpha_i=1\right\},$$

we get $S \subseteq C_1 \subseteq C_2 \subseteq \cdots$ “ $\to$ ”conv(S), or $C_{n+1} = \text{conv}(S)$. Our goal of the rest part of this section is to prove Carathéodory theorem.

Remark 1.2.2. There are some consequences of Carathéodory theorem. First, $\text{aff}(\text{conv}(S)) = \text{aff}(S)$. For the proof, showing ‘$\subseteq$’ part is enough. It is clear because *any affine combination of convex combination is indeed an affine combination.* Precisely, let $x \in \text{aff}(\text{conv}(S))$, and we get

$$x = \sum_{i=1}^{k} \alpha_i x_i \text{ for some } k \text{ and } x_i \in \text{conv}(S),\ \forall \alpha_i \geq 0$$

where

$$x_i = \sum_{j=1}^{k_i} \beta_{ij} y_{ij} \text{ for some } k_i \text{ and } y_{ij} \in S,\ \forall \beta_{ij} \geq 0,\ \sum_{j=1}^{k_i} \beta_{ij} = 1.$$

It implies that

$$x = \sum_{i=1}^{k} \sum_{j=1}^{k_i} \alpha_i \beta_{ij} y_{ij},\ \forall \alpha_i \beta_{ij} \geq 0,$$

and hence $x \in \text{aff}(S)$.

With this fact, we can define a dimension of convex hull, or a convex set, as $\dim(\text{conv}(S)) = \dim(\text{aff}(\text{conv}(S))) = \dim(\text{aff}(S))$. Or, for any convex set C, we can define $\dim(C) = \dim(\text{aff}(C))$. This definition coincides to our intuition, e.g., disk $\{(x, y) : x^2 + y^2 \leq 1\}$ on the plane has dimension 2. Furthermore, we get $\text{aff}(S) = \text{aff}(\text{cl}(S))$, where $\text{cl}(X)$ denotes the closure of X (It is clear because affine space is closed).

Example 1.2.3 (Convex hulls)**.** (a) $\text{conv}(\{x_1, x_2, \cdots, x_m\}) = \left\{ \sum_{i=1}^{m} \alpha_i x_i : \forall \alpha_i \geq 0, \sum_{i=1}^{m} \alpha_i = 1 \right\}$.

(b) If $C_1, C_2, \cdots, C_m$ are convex sets and $S = \bigcup_{i=1}^{m} C_i$, then

$$\text{conv}(S) = \left\{ \sum_{i=1}^{m} \alpha_i x_i : x_i \in C_i, \forall \alpha_i \geq 0, \sum_{i=1}^{m} \alpha_i = 1 \right\}.$$

(c) Let $A \in \mathbb{R}^{m \times n}$. Then

$$\text{conv}(A \cdot S) = A \cdot \text{conv}(S).$$

Proof. (b) Note that

$$\text{conv}(S) = \left\{ \sum_{i=1}^{k} \alpha_i x_i : x_i \in S,\ \forall \alpha_i \geq 0,\ \sum_{i=1}^{k} \alpha_i = 1 \text{ for some } k \right\}.$$

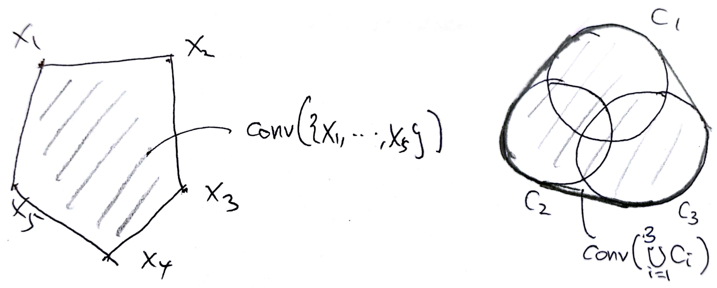


Figure 1.10: (a) and (b)

Let $x \in \text{conv}(S)$, and we get

$$x = \sum_{i=1}^{k} \alpha_i x_i \text{ for some } x_i \in S,\ \forall \alpha_i \geq 0,\ \sum_{i=1}^{k} \alpha_i = 1.$$

For any x_j, there is $C_{\pi(j)}$ for some $\pi(j) \in \{1, 2, \cdots, m\}$ such that $x_j \in C_{\pi(j)}$. (If there are many such C_k's, choose $\pi(j)$ be the smallest one, so that $\pi(j)$ can be well-defined) Then with rearrangement of index

$$\begin{aligned} x = \sum_{i=1}^{k} \alpha_i x_i &= \sum_{l=1}^{m} \sum_{\pi(j)=l} \alpha_j x_j \\ &= \sum_{l=1}^{m} \underbrace{\left(\sum_{\pi(i)=l} \alpha_i \right)}_{=:\beta_l} \underbrace{\sum_{\pi(j)=l} \frac{\alpha_j}{\sum_{\pi(i)=l} \alpha_i} x_j}_{=:y_l \in C_l} \\ &= \sum_{l=1}^{m} \beta_l y_l \end{aligned}$$

holds, which implies

$$x = \sum_{l=1}^{m} \beta_l y_l \in \left\{ \sum_{i=1}^{m} \alpha_i x_i : x_i \in C_i, \forall \alpha_i \geq 0, \sum_{i=1}^{m} \alpha_i = 1 \right\},$$

and '$\subseteq$' part is shown. '$\supseteq$' is trivial.

(c) Note that for $A \in \mathbb{R}^{m \times n}$,

If C is convex, $A \cdot C$ is convex,

and if C' is convex, $A^{-1}(C')$ is convex.

Thus, $A \cdot \text{conv}(S)$ is a convex set, containing $A \cdot S$, and we get $\text{conv}(A \cdot S) \subseteq A \cdot \text{conv}(S)$. Conversely, let

$$x \in \text{conv}(S), \text{ and represent } x = \sum_{i=1}^{k} \alpha_i x_i \text{ for some } k \text{ and } x_i \in S, \forall \alpha_i \geq 0, \sum_{i=1}^{k} \alpha_i = 1.$$

Then each Ax_i belongs to $A \cdot S$, and so $Ax = \sum_{i=1}^{k} \alpha_i \cdot Ax_i \in \text{conv}(A \cdot S)$, which implies $\text{conv}(A \cdot S) \supseteq A \cdot \text{conv}(S)$. □

Remark 1.2.4. Using the definition of convex hull, we can define "convexification" of non-convex function. Note that a function is convex iff its epigraph is. If we find a function whose epigraph is a convex hull of epigraph of given function, we can find "the nearest convex function" with given one. See figure 1.11.

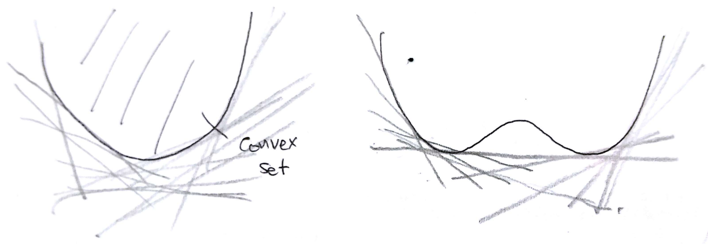


Figure 1.11: Convexification of non-convex function.

Now, we introduce a "conic hull."

Definition 1.2.5. $\sum_{i=1}^{k} \alpha_i x_i$ *is called a* ***conic combination*** *(or nonnegative combination) if* $\forall \alpha_i \geq 0$.

Note that comparing to convex combination, the condition $\sum \alpha_i = 1$ is omitted.

Proposition 1.2.6. *X is a convex cone if and only if*

$$X = \left\{ \sum_{i=1}^{k} \alpha_i x_i, \ x_i \in X, \ \forall \alpha_i \geq 0, \exists \alpha_i > 0, \ i = 1, 2, \cdots, k \textit{ for some } k \right\}.$$

Note that it is equivalent to

$$X = \left\{ \sum_{i=1}^{k} \alpha_i x_i, \ x_i \in X, \ \forall \alpha_i \geq 0, \sum_{i=1}^{k} \alpha_i > 0, \ i = 1, 2, \cdots, k \textit{ for some } k \right\}.$$

Recall that cone need not be convex. Also recall that even if convex cone cannot contain the origin. For convenience, we would consider convex cones containing 0.

Definition 1.2.7 (Conic hull)**.** ***Conic hull*** *of S is "defined" as*

$$cone(S) = \left\{ \sum_{i=1}^{k} \alpha_i x_i : x_i \in S,\ \alpha_i{\geq}0, \exists \alpha_i > 0,\ i = 1, 2, \cdots, k \text{ for some } k \right\}.$$

We defined $cone(S)$ as the smallest "convex cone" that contains S and 0.

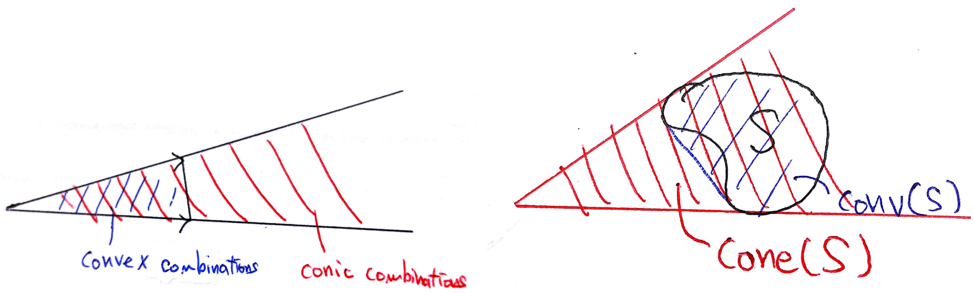


Figure 1.12: Conic combination and conic hull

Remark 1.2.8. (1) $0 \in \text{cone}(S)$, by definition of conic hull.

(2) Also, by definition of conic hull, $\text{cone}(S)$ is a convex cone.

(3)

$$\text{cone}(S) = \bigcup_{\lambda \geq 0} \lambda \cdot \text{conv}(S) = \text{conv}\left(\bigcup_{\lambda \geq 0} \lambda \cdot S \right)$$

holds, which implies $\text{cone}(S) \supseteq \text{conv}(S)$. It can be shown as following. First,

$$x \in \text{cone}(S) \Leftrightarrow x = \sum_{i=1}^{k} \alpha_i x_i \ \forall \alpha_i \geq 0, x_i \in S \Leftrightarrow x = \sum_{j=1}^{k} \alpha_j \sum_{i=1}^{k} \frac{\alpha_i}{\sum_j \alpha_j} x_i$$

$$\Leftrightarrow x \in \sum_{j=1}^{k} \alpha_j \cdot \text{conv}(S) \subseteq \bigcup_{\lambda \geq 0} \lambda \cdot \text{conv}(S).$$

Next,

$$x \in \text{cone}(S) \Rightarrow x = \sum_{i=1}^{k} \frac{\alpha_i}{\sum_j \alpha_j} \left(\sum_{j=1}^{k} \alpha_j x_i \right) \in \text{conv}\left(\bigcup_{\lambda \geq 0} \lambda \cdot S \right)$$

and

$$x \in \text{conv}\left(\bigcup_{\lambda \geq 0} \lambda \cdot S \right) \Rightarrow x = \sum_{i=1}^{k} \alpha_i \lambda_i x_i \ \forall \alpha_i \geq 0, \sum \alpha_i = 1,\ x_i \in S \in \text{cone}(S)$$

yields the result.

(4)

$$\text{cone}(S) = \bigcap (\text{convex cone} \supseteq S) \cup \{0\}.$$

(5) Remark that, cone(S) need not be closed, even if S is compact. For example, let $S = \{(x_1, x_2) : x_1^2 + (x_2 - 1)^2 \le 1\}$. Then S is compact, but

$$\text{cone}(S) = \{(x_1, x_2) : x_2 > 0\} \cup \{0\}$$

is not closed.

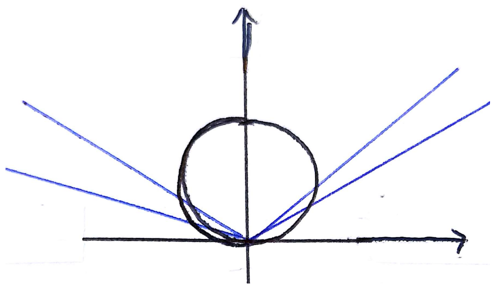


Figure 1.13: Example in (5)

Now we are ready to prove Carathéodory theorem. Indeed, it also says similar argument about conic hull.

Theorem 1.2.9 (Carathéodory). *Let $\phi \neq S \subseteq \mathbb{R}^n$ be a nonempty set.*

(a) For any $0 \neq x \in$ cone(S) can be represented as a "positive combination" of linearly independent points (vectors) from S, i.e.,

$$x = \sum_{i=1}^{k} \alpha_i x_i, \;\; x_i \in S, \;\; \alpha_i > 0, \;\; i = 1, 2, \cdots, k$$

for some k, where $x_1, \cdots, x_k$ are linearly independent vectors.

(b) For any $x \in$ conv(S) can be represented as a convex combination of at most $n+1$ points from S.

Proof. (a) We can find a conic combination representation. Leaving zero coefficients out, we can find "the smallest integer" m such that

$$x = \sum_{i=1}^{m} \alpha_i x_i, \text{ where } \alpha_i > 0,\ x_i \in S,\ i = 1, 2, \cdots, m.$$

If x_i's are linearly dependent, then $\exists \lambda_1, \cdots, \lambda_m$ s.t. $\sum_{i=1}^{m} \lambda_i x_i = 0,\ (\lambda_1, \cdots, \lambda_m) \neq 0$, and at least one $\lambda_i > 0$. (If all of λ_i's are negative, consider $-\lambda_i$ instead.) Let $\mathcal{I} = \{i : \lambda_i > 0\}$. Choose k such that

$$\gamma := \frac{\alpha_k}{\lambda_k} = \min\left\{\frac{\alpha_i}{\lambda_i} : i \in \mathcal{I}\right\},$$

and then for $\gamma > 0$, we get

$$x = \sum_{i=1}^{m} \alpha_i x_i - \gamma \underbrace{\sum_{i=1}^{m} \lambda_i x_i}_{=0} = \sum_{i=1}^{m} \underbrace{(\alpha_i - \gamma\lambda_i)}_{\geq 0} x_i,$$

where $\alpha_k - \gamma\lambda_k = 0$. It is contradictory to the assumption that m is minimal.

(b) Let $x \in \text{conv}(X)$. Then

$$x = \sum_{j=1}^{m} \gamma_j x_j \text{ for some } \gamma_j > 0, x_j \in X, \sum_{i=1}^{m} \gamma_j = 1.$$

(Remark that we only considered *positive combination*, i.e., $\gamma_i > 0$. It is always possible because otherwise we can just omit the zero coefficients.) Not define

$$Y = \{(y, 1) : y \in X\}.$$

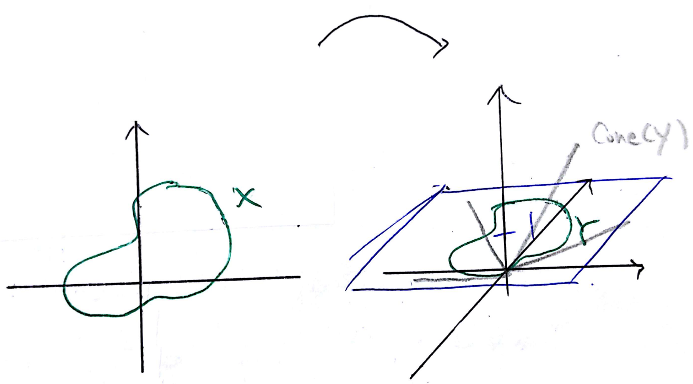


Figure 1.14: Proof of Carathéodory theorem.

Then from

$$\begin{pmatrix} x \\ 1 \end{pmatrix} = \sum_{j=1}^{m} \gamma_j \begin{pmatrix} x_j \\ 1 \end{pmatrix},$$

we get $\begin{pmatrix} x \\ 1 \end{pmatrix} \in \text{cone}(Y)$. Thus by (a), we can find $\begin{pmatrix} x_i \\ 1 \end{pmatrix} \in Y$ and $\alpha_i > 0$ s.t.

$$\begin{pmatrix} x \\ 1 \end{pmatrix} = \sum_{i=1}^{k} \alpha_i \begin{pmatrix} x_i \\ 1 \end{pmatrix},$$

where $\alpha_i > 0$ and $\left\{ \begin{pmatrix} x_1 \\ 1 \end{pmatrix}, \begin{pmatrix} x_2 \\ 1 \end{pmatrix}, \cdots, \begin{pmatrix} x_k \\ 1 \end{pmatrix} \right\} \subseteq \mathbb{R}^{n+1}$ are linearly independent. It implies that

$$x = \sum_{i=1}^{k} \alpha_i x_i,\ \ \alpha_i > 0,\ \sum_{i=1}^{k} \alpha_i = 1$$

and $k \le n+1$ from independence of k vectors in $\mathbb{R}^{n+1}$. □

Remark 1.2.10. (i) For (b), $x_1, \cdots, x_k$'s need not be linearly independent. Just independence of $\begin{pmatrix} x_1 \\ 1 \end{pmatrix}, \cdots, \begin{pmatrix} x_k \\ 1 \end{pmatrix}$ should be! Augmenting new coordinate is a technique used frequently.

(ii) Choice of x_i's is not unique; x_i's can be chosen s.t. $\{x_2 - x_1, \cdots, x_m - x_1\}$ are linearly independent. In this view, we think "x_1 as the new origin."

Following is our last result in this section.

Proposition 1.2.11. *Convex hull of compact set is also compact.*

Proof. Let $X \subseteq \mathbb{R}^n$ be a compact set, and suppose that $X \neq \phi$ (If $X = \phi$, there is nothing to prove). Let $\{x^{(k)}\}_{k=1}^{\infty} \subseteq \text{conv}(X)$ be an arbitrary sequence. By Carathéodory theorem,

$$x^{(k)} = \sum_{i=1}^{n+1} \alpha_i^{(k)} x_i^{(k)} \text{ for some } x_i^{(k)} \in X,\ \alpha_i^{(k)} \ge 0 \text{ and } \sum_{i=1}^{n+1} \alpha_i^{(k)} = 1.$$

Consider a new vector

$$(\alpha_1^{(k)}, \cdots, a_{n+1}^{(k)}, x_1^{(k)}, \cdots, x_{n+1}^{(k)})^\top$$

and sequence of these vectors

$$\{(\alpha_1^{(k)}, \cdots, a_{n+1}^{(k)}, x_1^{(k)}, \cdots, x_{n+1}^{(k)})^\top\}_{k=1}^{\infty} \subseteq \mathbb{R}^{n+1} \times X^{n+1}.$$

By boundedness of $\alpha_i^{(k)}$ and compactness of X, such sequence is bounded, so by Bolzano-Weierstrass theorem, there is a convergent subsequence

$$\{(\alpha_1^{(k')}, \cdots, a_{n+1}^{(k')}, x_1^{(k')}, \cdots, x_{n+1}^{(k')})^\top\} \subseteq \{(\alpha_1^{(k)}, \cdots, a_{n+1}^{(k)}, x_1^{(k)}, \cdots, x_{n+1}^{(k)})^\top\}$$

converging to $(\alpha_1, \cdots, \alpha_{n+1}, x_1, \cdots, x_{n+1})^\top$, and it should satisfy

$$\sum_{i=1}^{n+1} \alpha_i = 1, \ \alpha_i \geq 0, \text{ and } x_i \in X.$$

Then

$$x^{(k')} = \sum_{i=1}^{n+1} \alpha_i^{(k')} x_i^{(k')} \xrightarrow[k\to\infty]{} \sum_{i=1}^{n+1} \alpha_i x_i \in \mathrm{conv}(X)$$

holds. Thus, every sequence in conv(X) has a convergent subsequence whose limit point is in conv(X), so conv(X) is closed. Note that conv(X) is bounded by boundedness of X. Therefore, by Heine-Borel, conv(X) is compact. □

Remark 1.2.12. Note that if X is not compact, even if it is closed, conv(X) even need not be closed. For example, let $X = \{(0,0)\} \cup \{(x,y) : xy \geq 1, \ x, y \geq 0\}$. Then X is closed, but its convex hull conv(X) $= \{(0,0)\} \cup \{(x,y) : x, y > 0\}$ is not closed.

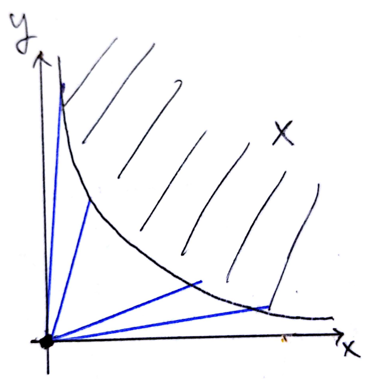


Figure 1.15: The set $X = \{(0,0)\} \cup \{(x,y) : xy \geq 1, \ x, y \geq 0\}$ and its convex hull.

1.3 Relative interior and closure

In this section, we study some generic topological properties of convex sets and functions. Note that an interior of a convex set may be empty set. However, if we consider a *relative interior* with respect to an affine hull, then we can construct some topology on convex sets. From now one, for convenience, let $C \subseteq \mathbb{R}^n$ be a nonempty convex set, and $\mathrm{cl}(C)$ and $\mathrm{int}(C)$ be its closure and interior, respectively.

Definition 1.3.1 (Relative interior)**.** *x is a* ***relative interior point*** *of C if $x \in C$ and $\exists \epsilon > 0$ s.t. for an open ball $B(x, \epsilon)$,*

$$B(x, \epsilon) \cap \mathrm{aff}(C) \subseteq C.$$

We also refer to x as an "interior point of C relative to $\mathrm{aff}(C)$." Relative interior $\mathrm{ri}(C)$ (or relint(C)) is a collection of relative interior points. A set C is said to be ***relative open*** *if $C = \mathrm{ri}(C)$.*

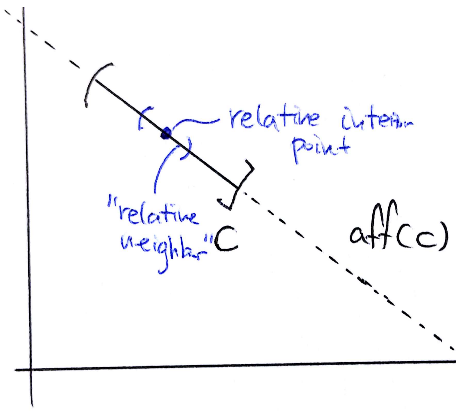


Figure 1.16: Relative interior

Definition 1.3.2 (Relative boundary)**.** *$x \in C$ is called* ***relative boundary point*** *if $x \in \mathrm{cl}(C)$ but $x \notin \mathrm{ri}(C)$. The set of relative boundary points is denoted as $\mathrm{rb}(C)$.*

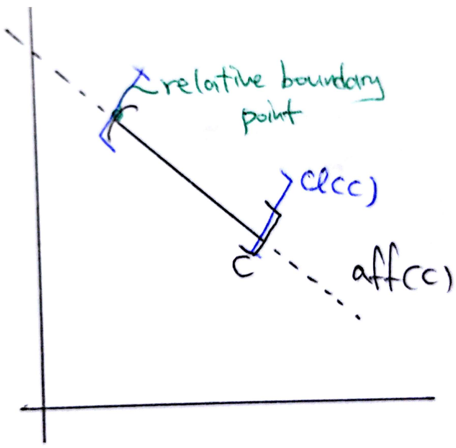


Figure 1.17: Relative boundary

Example 1.3.3. If C is affine, $\mathrm{ri}(C) = C$ and $\mathrm{rb}(C) = \phi$.

Following is the first result of relative topology, which is very useful.

Proposition 1.3.4 (Line Segment Principle)**.** *Let $x \in \text{ri}(C)$ and $\bar{x} \in \text{cl}(C)$. Then for any point on the "line segment" connecting x and $\bar{x}$ is contained in* ri(C)*, except possibly $\bar{x}$.*

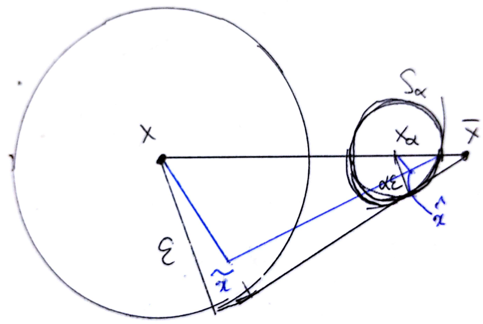


Figure 1.18: Line Segment Principle

Proof. i) $\bar{x} \in C$.

From $x \in \text{ri}(C)$, $\exists \epsilon > 0$ s.t. $B(x, \epsilon) \cap \text{aff}(C) \subseteq C$. Let $x_\alpha = \alpha x + (1-\alpha)\bar{x}$ be a point on the line segment. We may assume $0 < \alpha \le 1$. Define $S_\alpha = B(x_\alpha, \alpha\epsilon)$. Then for any $\hat{x} \in S_\alpha \cap \text{aff}(C)$, $\exists \beta \in (0, 1]$ and $\exists \tilde{x} \in B(x, \epsilon) \cap \text{aff}(C)$ s.t.

$$\hat{x} = \beta\tilde{x} + (1-\beta)\bar{x}.$$

Note that, we are considering relative topology, so we get $\hat{x}$ from $S_\alpha \cap$"aff(C)". (Remark: Indeed, for proper $\hat{x}$, $\beta = \alpha$ in here. Choose

$$\tilde{x} := x - \frac{x_\alpha - \hat{x}}{\alpha}.$$

Then from $\hat{x} \in S_\alpha$, $\tilde{x} \in B(x, \epsilon)$, and $\hat{x} = \alpha\tilde{x} + (1-\alpha)\bar{x}$.) Note that $\tilde{x} \in B(x, \epsilon) \cap \text{aff}(C) \subseteq C$, and from the assumption that $\bar{x} \in C$, we get $\hat{x} \in C$. Since $\hat{x}$ was arbitrary, we get $S_\alpha \cap \text{aff}(C) \subseteq C$, which implies that x_α is a relative interior point.

ii) $\bar{x} \notin C$.

Use limit argument. There is a sequence $\{x_k\} \subseteq C$ s.t. $x_k \xrightarrow[k\to\infty]{} x$. Let $x_{k,\alpha} = \alpha x + (1-\alpha)x_k$, and then by i), $x_{k,\alpha} \in \text{ri}(C)$ for any k. Letting $k \to \infty$, $x_{k,\alpha} \to x_\alpha$, so for sufficiently large k,

$$\|x_{k,\alpha} - x_\alpha\| < \frac{\alpha\epsilon}{2},$$

and hence

$$B\left(x_\alpha, \frac{\alpha\epsilon}{2}\right) \subseteq B(x_{k,\alpha}, \alpha\epsilon).$$

($\because \|y - x_\alpha\| < \alpha\epsilon/2 \Rightarrow \|y - x_{k,\alpha}\| \le \|y - x_\alpha\| + \|x_\alpha - x_{k,\alpha}\| < \alpha\epsilon$) Therefore, "relative neighbor" of x_α satisfies

$$B\left(x_\alpha, \frac{\alpha\epsilon}{2}\right) \cap \mathrm{aff}(C) \subseteq B(x_{k,\alpha}, \alpha\epsilon) \cap \mathrm{aff}(C) \subseteq C,$$

which implies $x_\alpha \in \mathrm{ri}(C)$. □

Proposition 1.3.5 (Nonemptiness of relative interior)**.** *Relative interior of every nonempty convex set is nonempty. In precise, for nonempty convex set C,*

(a) $\mathrm{ri}(C) \neq \phi$*, and* $\mathrm{ri}(C)$ *is also convex. Further,* $\mathrm{aff}(\mathrm{ri}(C)) = \mathrm{aff}(C)$*.*

(b) If $m = \dim(\mathrm{aff}(C)) > 0$ *("only meaningful case"),* $\exists x_0, x_1, \cdots, x_m \in \mathrm{ri}(C)$ *s.t.* $\mathit{span}(x_1 - x_0, \cdots, x_m - x_0) // \mathrm{aff}(C)$*. ("relative basis regarding one point as the origin")*

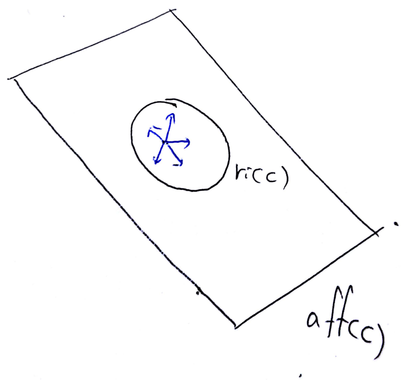


Figure 1.19: Proposition 1.3.5(b).

Proof. (a) First, convexity of $\mathrm{ri}(C)$ is obvious from line segment principle.

For nonemptiness, WLOG assume $0 \in C$ (Otherwise, shift). Then $\mathrm{aff}(C)$ becomes a subspace. Let $m := \dim(\mathrm{aff}(C))$.

<u>Case 1.</u> $m = 0$. Then $\mathrm{aff}(C) = \{0\}$ and so $C = \{0\}$. In this case, clearly $\mathrm{ri}(C) = \{0\}$ is nonempty.

<u>Case 2.</u> $m > 0$. By Carathéodory theorem, C contains m linearly independent $z_1, \cdots, z_m$ (subspace is also a cone). It implies that $\mathrm{span}(z_1, \cdots, z_m) = \mathrm{aff}(C)$. Then $\{z_1, \cdots, z_m\}$ becomes a basis, and we can define

$$X := \left\{ \sum_{i=1}^{m} \alpha_i z_i : \alpha_i > 0,\ i = 1, 2, \cdots, m,\ \sum_{i=1}^{m} \alpha_i < 1 \right\}.$$

Note that X is nonempty, for instance, $\dfrac{z_1 + \cdots + z_m}{m+1} \in X$. Further, for any $x \in X$,

$$x = \sum_{i=1}^{m} \alpha_i z_i = \sum_{i=1}^{m} \alpha_i z_i + \alpha_{m+1} \cdot 0 \in C$$

for $\alpha_{m+1} = 1 - (\alpha_1 + \cdots + \alpha_m)$, and hence $X \subseteq C$. Now our claim is that X is open relative to $\text{aff}(C)$.

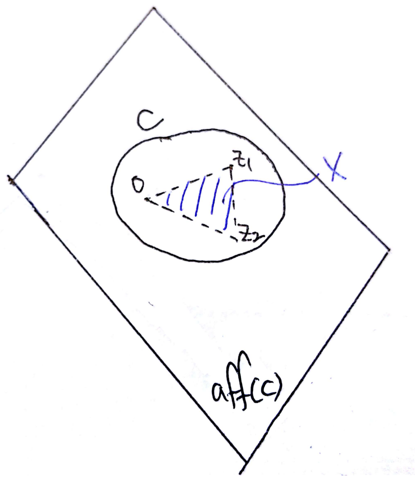


Figure 1.20: Proof of Proposition 1.3.5(a): C and X.

(Motivation: Because $\text{aff}(C)$ is the space similar as $\mathbb{R}^m$, we would think $\alpha_i \in \mathbb{R}^m$ rather than $z_i \in \mathbb{R}^n$!) Let $\bar{x} \in X \subseteq C$. We will show that $\exists \epsilon > 0$ s.t. $B(\bar{x}, \epsilon) \cap \text{aff}(C) \subseteq X$. Let $x \in \text{aff}(C)$. Then since $\{z_1, \cdots, z_m\}$ is a basis, for $Z = [z_1, z_2, \cdots, z_m] \in \mathbb{R}^{n \times m}$, there are coefficient vectors $\bar{\alpha}$ and α such that

$$\bar{x} = Z\bar{\alpha}, \ x = Z\alpha.$$

Note that from $\bar{x} \in X$ we get $\bar{\alpha}_i > 0$ and $\sum_{i=1}^m \bar{\alpha}_i < 1$. Then we get

$$\|x - \bar{x}\|^2 = (\alpha - \bar{\alpha})^\top Z^\top Z(\alpha - \bar{\alpha}).$$

Let λ_{min} be minimum eigenvalue of $Z^\top Z$. Then by positive definiteness of $Z^\top Z$, $\lambda_{min} > 0$, so $\gamma^2 = \lambda_{min}$ for some $\gamma > 0$. Then by spectral theorem,

$$\|x - \bar{x}\|^2 = (\alpha - \bar{\alpha})^\top Z^\top Z(\alpha - \bar{\alpha}) \geq \gamma^2 \|\alpha - \bar{\alpha}\|^2.$$

Now let

$$A := \left\{ (\alpha_1, \cdots, \alpha_m)^\top : \sum_{i=1}^m \alpha_i < 1,\ \alpha_i > 0 \right\} \subseteq \mathbb{R}^m.$$

By definition, $\alpha \in A \Leftrightarrow Z\alpha \in X$, and it can be easily shown that A is an open set. Therefore, $\exists \epsilon > 0$ s.t.

$$B\left(\bar{\alpha}, \frac{\epsilon}{\gamma}\right) \subseteq A$$

from $\bar{\alpha} \in A$, and hence if $x \in B(\bar{x}, \epsilon)$,

$$\|\alpha - \bar{\alpha}\|^2 \le \frac{1}{\gamma^2}\|x - \bar{x}\|^2 < \epsilon^2,$$

which implies

$$\alpha \in B\left(\bar{\alpha}, \frac{\epsilon}{\gamma}\right) \subseteq A,$$

i.e., $x \in X$. In summary, $x \in B(\bar{x}, \epsilon) \cap \mathrm{aff}(C) \Rightarrow x \in X$ holds. It means that $B(\bar{x}, \epsilon) \cap \mathrm{aff}(C) \subseteq X \subseteq C$. Especially, from $\bar{x} \in X$, $\bar{x} \in C$, so $\bar{x}$ becomes a relative interior point of C. Therefore, $\bar{x} \in \mathrm{ri}(C)$, i.e., $\mathrm{ri}(C) \neq \phi$. ⊠

Finally, we will show $\mathrm{aff}(\mathrm{ri}(C)) = \mathrm{aff}(C)$. First, note that

$$\mathrm{cl}(X) = \left\{ \sum_{i=1}^m \alpha_i z_i : \alpha_i \ge 0,\ i = 1, 2, \cdots, m,\ \sum_{i=1}^m \alpha_i \le 1 \right\}.$$

Then $0 \in \mathrm{cl}(X)$ and $z_i \in \mathrm{cl}(X)$, so $\mathrm{aff}(\mathrm{cl}(X)) \supseteq \mathrm{span}(z_1, \cdots, z_m)$. Thus,

$$\mathrm{span}(z_1, \cdots, z_m) \subseteq \mathrm{aff}(\mathrm{cl}(X)) = \mathrm{aff}(X) \subseteq \mathrm{aff}(C) = \mathrm{span}(z_1, \cdots, z_m)$$

implies that $\mathrm{aff}(X) = \mathrm{aff}(C)$. Moreover, in previous part, actually we showed that any point of X is relative interior point of C, i.e.,

$$X \subseteq \mathrm{ri}(C).$$

Therefore,

$$\mathrm{aff}(X) \subseteq \mathrm{aff}(\mathrm{ri}(C)) \subseteq \mathrm{aff}(C) = \mathrm{aff}(X)$$

implies $\mathrm{aff}(\mathrm{ri}(C)) = \mathrm{aff}(C)$.

(b) By (a), $\exists x_0 \in \mathrm{ri}(C)$. Translate C to $C - x_0$ in order to be $0 \in C - x_0$. Note that $\mathrm{ri}(C - x_0) = \mathrm{ri}(C) - x_0$ holds, which implies that $0 \in \mathrm{ri}(C - x_0)$. Thus by Carathéodory again,

$$\exists z_1, z_2, \cdots, z_m \in C - x_0 \text{ s.t. } z_1, \cdots, z_m \text{ are linearly independent},$$

i.e.,

$$\mathrm{span}(z_1, z_2, \cdots, z_m) = \mathrm{aff}(C - x_0).$$

Let $\alpha \in (0,1)$. Then by line segment principle,

$$(1-\alpha)\cdot \underbrace{0}_{\in \mathrm{ri}(C-x_0)} + \alpha \cdot \underbrace{z_i}_{\in C-x_0} \in \mathrm{ri}(C - x_0).$$

Let $x_i = x_0 + \alpha z_i$. Then $x_1, x_2, \cdots, x_m \in \mathrm{ri}(C)$ and $\mathrm{span}(x_1 - x_0, \cdots, x_m - x_0) = \mathrm{aff}(C - x_0)$, which is the desired result. $\square$

Proposition 1.3.6 (Prolongation Lemma)**.** *Let $C \neq \phi$ be a nonempty convex set. Then*

$$x \in \mathrm{ri}(C) \Leftrightarrow \forall \bar{x} \in C\ \exists \gamma > 0\ s.t.\ x + \gamma(x - \bar{x}) \in C.$$

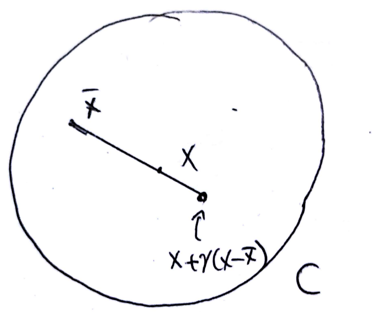


Figure 1.21: Prolongation lemma.

Proof. $\Rightarrow$) From $x \in \mathrm{ri}(C)$, $\exists \epsilon > 0$ s.t. $B(x,\epsilon) \cap \mathrm{aff}(C) \subseteq C$. Then $\forall \bar{x} \in C$, letting

$$\gamma = \frac{\epsilon}{2\|x - \bar{x}\|},$$

we get $x + \gamma(x - \bar{x}) \in B(x, \epsilon) \cap \mathrm{aff}(C) \subseteq C$.

$\Leftarrow$) In particular, for some $\bar{x} \in \mathrm{ri}(C)$ (by proposition 1.3.5, we can choose such $\bar{x}$), the condition should be satisfied. If $x \neq \bar{x}$, there is nothing to prove, so assume that $x \neq \bar{x}$. Then by the condition,

$$\exists y = x + \gamma(x - \bar{x}) \in C.$$

Then by line-segment principle, ($\bar{x} \in \mathrm{ri}(C)$)

$$x = \frac{\gamma}{1+\gamma}\bar{x} + \frac{1}{1+\gamma}y \in \mathrm{ri}(C).$$

Remark 1.3.7. Note that line segment principle tells the thing about the point of *internal* division, while prolongation lemma says that there exists at least one point of *external* division in the convex set. It is very important to understand that

z is the point of internal division of x and y $\Rightarrow$ y is the point of external division of x and z.

In other words, we proved prolongation lemma using line segment principle above, with the relationship between internal and external divisions. However, '$\Leftarrow$' does not hold in general, so it's hard to say that two statements (line segment principle and prolongation lemma) are equivalent.

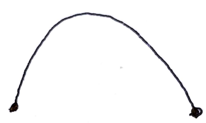
There is an important result of prolongation lemma. It is about "minimization of concave functions."

Proposition 1.3.8. *Let $\phi \neq X \subseteq \mathbb{R}^n$ be a convex set, and $f : X \to \mathbb{R}$ be a concave function. Let*

$$X^* := \left\{ x^* \in X : f(x^*) = \inf_{x \in X} f(x) \right\} = \arg\min_{x \in X} f(x).$$

If $X^* \cap \mathrm{ri}(X) \neq \phi$, then $X^* = X$, i.e., f is a constant function on X.

Remark 1.3.9. It means that, minimum point of concave function should be in the (relative) boundary.



Proof. From nonemptiness, we can choose $x^* \in X^* \cap \mathrm{ri}(X)$. Also let $x \in X$. Then by prolongation lemma, $\exists \gamma > 0$ s.t.

$$\hat{x} = x^* + \gamma(x^* - x) \in X.$$

Then by concavity,

$$f(x^*) = f\left(\frac{1}{1+\gamma}\hat{x} + \frac{\gamma}{1+\gamma}x\right) \geq \frac{1}{1+\gamma}f(\hat{x}) + \frac{\gamma}{1+\gamma}f(x)$$

holds, and from $x^* \in X$,

$$f(\hat{x}) \geq f(x^*), f(x) \geq f(x^*),$$

which imply

$$f(x^*) \geq \frac{1}{1+\gamma} f(\hat{x}) + \frac{\gamma}{1+\gamma} f(x) \geq \frac{1}{1+\gamma} f(x^*) + \frac{\gamma}{1+\gamma} f(x^*) = f(x^*).$$

Therefore, we get

$$\frac{1}{1+\gamma} f(\hat{x}) + \frac{\gamma}{1+\gamma} f(x) = f(x^*),$$

i.e.,

$$f(x) = f(\hat{x}) = f(x^*).$$

Since x was arbitrary, we get $X^* = X$. $\square$

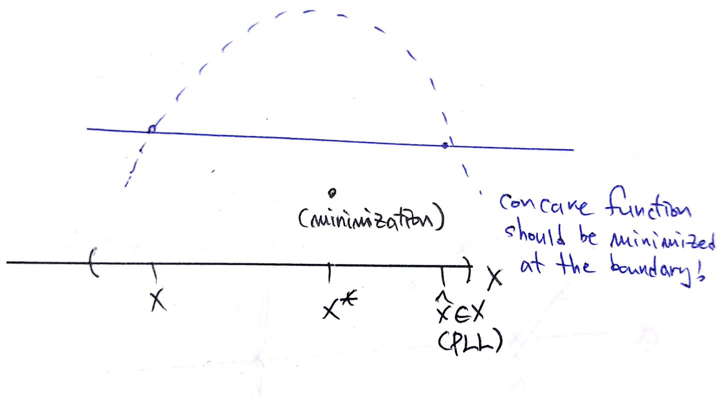


Figure 1.22: Proof of proposition 1.3.8.

1.3.1 Calculus of relative interior and closure

From now on, we see intersection, sum, linear transformation, etc... of relative interior or closure, and their topological properties.

Proposition 1.3.10. *Let $\phi \neq C$ be a nonempty convex set. Then*

(a) $\text{cl}(C) = \text{cl}(\text{ri}(C))$.

(b) $\text{ri}(C) = \text{ri}(\text{cl}(C))$.

(c) If $\bar{C} \neq \phi$ is another nonempty convex set, then

$$\text{ri}(C) = \text{ri}(\bar{C}) \Leftrightarrow \text{cl}(C) = \text{cl}(\bar{C}) \Leftrightarrow \text{ri}(C) \subseteq \bar{C} \subseteq \text{cl}(C).$$

Sketch of Proof. (c) is clear by (a) and (b). For (a), let $x \in \mathrm{cl}(C)$. Then $\exists\{x_n\} \subseteq C$ s.t. $x_n \to x$. For fixed $\tilde{x} \in \mathrm{ri}(C)$, and $\alpha_n \searrow 0$, by line segment principle, $\alpha_n \tilde{x} + (1-\alpha_n)x_n \in \mathrm{ri}(C)$, and it converges to x. Therefore $x \in \mathrm{cl}(\mathrm{ri}(C))$. For (b), let $x \in \mathrm{ri}(\mathrm{cl}(C))$. Then by prolongation lemma, for $y \in \mathrm{ri}(C)$, $\exists \gamma > 0$ s.t. $z = x + \gamma(x-y) \in \mathrm{cl}(C)$. Thus $x = (1+\gamma)^{-1}z + \gamma(1+\gamma)^{-1}y \in \mathrm{ri}(C)$, by line segment principle. □

Proposition 1.3.11. *Let $\phi \neq C \subseteq \mathbb{R}^n$ be a nonempty convex set, and $A \in \mathbb{R}^{m\times n}$. Then*

(a) $A \cdot \mathrm{ri}(C) = \mathrm{ri}(A \cdot C)$.

(b) $A \cdot \mathrm{cl}(C) \subsetneq \mathrm{cl}(A \cdot C)$.

(c) If C is bounded, $A \cdot \mathrm{cl}(C) = \mathrm{cl}(A \cdot C)$.

Sketch of Proof. For (b), if $x \in A \cdot \mathrm{cl}(C)$, $\exists y \in \mathrm{cl}(C)$ and $y_n \in C$ s.t. $y_n \to y,\ Ay = x$, which implies $Ay_n \to Ay$, i.e., $Ay \in \mathrm{cl}(A\cdot C)$. Remark that: equalness does not hold, because *even if $\exists\{x_n\} \subseteq A \cdot C$ s.t. $Ax_n \to Ax \in \mathrm{cl}(A\cdot C)$, we cannot say that $x_n \to x$.* To solve this problem, we need boundedness, to guarantee the existence of convergent subsequence of $\{x_n\}$. For (a),

$$\mathrm{ri}(A\cdot\mathrm{ri}(C)) \subseteq A \cdot C \subseteq A \cdot \mathrm{ri}(C) \subseteq A \cdot \mathrm{cl}(C) = A \cdot \mathrm{cl}(\mathrm{ri}(C)) \subseteq \mathrm{cl}(A \cdot \mathrm{ri}(C))$$

implies $\mathrm{ri}(A \cdot C) = \mathrm{ri}(A \cdot \mathrm{ri}(C)) \subseteq A \cdot \mathrm{ri}(C)$. Conversely let $x \in A \cdot \mathrm{ri}(C)$. $\exists y \in \mathrm{ri}(C)$ s.t. $x = Ay$. By prolongation lemma, $\forall z \in C\ \exists \gamma > 0$ s.t. $y + \gamma(y-z) \in C$, which implies $\forall Az \in A\cdot C\ Ay + \gamma(Ay - Az) \in A \cdot C$, and hence by prolongation lemma again, $x = Ay \in \mathrm{ri}(A\cdot C)$. □

Example 1.3.12. In here we see a counter-example that $A \cdot \mathrm{cl}(C) = \mathrm{cl}(A\cdot C)$ does not hold. Consider

$$C = \{(x,y) : x > 0, y > 0, xy \geq 1\}$$

and $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. Then C is convex, closed, and

$$A \cdot C = \{(x,0) : x > 0\}.$$

Thus, we get

$$\mathrm{cl}(A\cdot C) = \{(x,0) : x \geq 0\},\ A \cdot \mathrm{cl}(C) = A \cdot C = \{(x,0) : x > 0\}.$$

Proposition 1.3.13. *Let C_1, C_2 be nonempty convex sets. Then,* *(a)* $\mathrm{ri}(C_1 + C_2) = \mathrm{ri}(C_1) + \mathrm{ri}(C_2)$.

(b) $\mathrm{cl}(C_1 + C_2) \supsetneq \mathrm{cl}(C_1) + \mathrm{cl}(C_2)$.

(c) If either C_1 or C_2 is bounded, then $\mathrm{cl}(C_1 + C_2) = \mathrm{cl}(C_1) + \mathrm{cl}(C_2)$.

Sketch of proof. (a) Let $x + y \in \mathrm{ri}(C_1) + \mathrm{ri}(C_2)$, where $x \in \mathrm{ri}(C_1)$ and $y \in \mathrm{ri}(C_2)$. By PLL, $\forall \tilde{x} \in C_1,\ \tilde{y} \in C_2\ \exists \gamma_1 > 0, \gamma_2 > 0$ s.t.

$$x + \gamma_1(x - \tilde{x}) \in C_1,\ y + \gamma_2(y - \tilde{y}) \in C_2.$$

WLOG $\gamma_1 > \gamma_2$. LSP implies $x + \gamma_2(x - \tilde{x}) \in C_1$ and so $(x+y) + \gamma_2((x+y) - (\tilde{x} + \tilde{y})) \in C_1 + C_2$, and hence PLL says '$\supseteq$.' '$\subseteq$' comes from

$$\mathrm{ri}(\mathrm{ri}(C_1)+\mathrm{ri}(C_2)) \subseteq \mathrm{ri}(C_1)+\mathrm{ri}(C_2) \subseteq C_1+C_2 \subseteq \mathrm{cl}(C_1)+\mathrm{cl}(C_2) = \mathrm{cl}(\mathrm{ri}(C_1))+\mathrm{cl}(\mathrm{ri}(C_2)) \underset{\text{(b)}}{\subseteq} \mathrm{cl}(\mathrm{ri}(C_1)+\mathrm{ri}(C_2)),$$

which implies that

$$\mathrm{ri}(\mathrm{ri}(C_1) + \mathrm{ri}(C_2)) = \mathrm{ri}(C_1 + C_2)$$

from proposition 1.3.10(c). Or, using $\mathrm{ri}(C_1 \times C_2) = \mathrm{ri}(C_1) \times \mathrm{ri}(C_2)$ and proposition 1.3.11(c) to $A : (x_1, x_2) \mapsto x_1 + x_2$ we can obtain the same result.

(b), (c) Let $x + y \in \mathrm{cl}(C_1) + \mathrm{cl}(C_2)$, where $x \in \mathrm{cl}(C_1)$ and $y \in \mathrm{cl}(C_2)$. Then $\exists x_n \in C_1,\ y_n \in C_2$ s.t. $x_n \to x,\ y_n \to y$. It implies $x_n + y_n \to x + y$, and from $x_n + y_n \in C_1 + C_2$ we get $x + y \in \mathrm{cl}(C_1 + C_2)$. However, we cannot guarantee that $x_n \to x$ and $y_n \to y$ hold from $x_n + y_n \to x + y$. Rather, with boundedness, we can say that such "subsequences" exist. □

Example 1.3.14. Again, we see a counter-example. Let

$$C_1 = \{(x, y) : x > 0, y > 0, xy \geq 1\} \text{ and } C_2 = \{(0, y) : y \in \mathbb{R}\}.$$

Then

$$C_1 + C_2 = \{(x, y) : x > 0\}$$

is an open half space, so $\mathrm{cl}(C_1 + C_2) = \{(x, y) : x \geq 0\}$, while $\mathrm{cl}(C_1) + \mathrm{cl}(C_2) = C_1 + C_2$.

Proposition 1.3.15. *Let C_1, C_2 be nonempty convex sets. Then,*

(a) $\mathrm{ri}(C_1 \cap C_2) \supsetneq \mathrm{ri}(C_1) \cap \mathrm{ri}(C_2)$.

(b) $\mathrm{cl}(C_1 \cap C_2) \subsetneq \mathrm{cl}(C_1) \cap \mathrm{cl}(C_2)$.

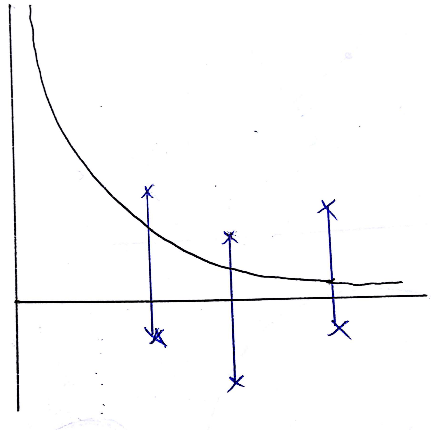


Figure 1.23: Example 1.3.14.

(c) If $\mathrm{ri}(C_1) \cap \mathrm{ri}(C_2) \neq \phi$*, then both "=" hold in (a) and (b).*

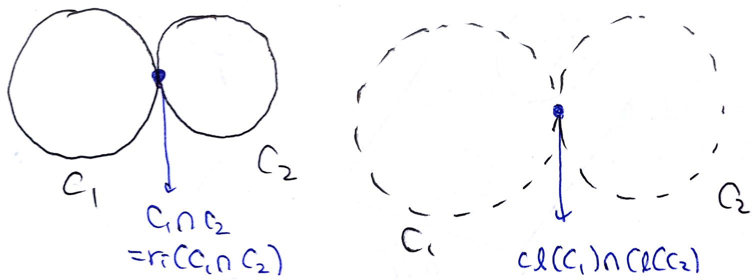


Figure 1.24: Proposition 1.3.15.

Sketch of proof. Let $x \in \mathrm{ri}(C_1) \cap \mathrm{ri}(C_2)$. Then $\forall y_1 \in C_1, \forall y_2 \in C_2$, $\exists \gamma > 0$ s.t

$$x + \gamma(x - y_1) \in C_1, \ x + \gamma(x - y_2) \in C_2.$$

(Choose such γ_1 and γ_2, and get γ be their minimum) Letting $y \in C_1 \cap C_2$, we get $x + \gamma(x - y) \in C_1 \cap C_2$, and hence $x \in \mathrm{ri}(C_1 \cap C_2)$. For converse, even though $x \in \mathrm{ri}(C_1 \cap C_2)$, we can only say that

$$\forall y \in C_1 \cap C_2 \ \exists \gamma > 0 \text{ s.t. } x + \gamma(x - y) \in C_1 \cap C_2,$$

but not '$\forall y \in C_1$' argument. However, if $\mathrm{ri}(C_1) \cap \mathrm{ri}(C_2) \neq \phi$, choose one element y from there, and then

$$z := x + \gamma(x - y) \in C_1 \cap C_2$$

implies

$$x = \frac{1}{1+\gamma}z + \frac{\gamma}{1+\gamma}y \in \mathrm{ri}(C_1)$$

by LSP. Remark that the assumption $y \in \mathrm{ri}(C_1)$ is critical to apply LSP!

Now let $x \in \mathrm{cl}(C_1 \cap C_2)$. Then $\exists x_n \in C_1 \cap C_2$ s.t. $x_n \to x$, which implies $x \in \mathrm{cl}(C_1) \cap \mathrm{cl}(C_2)$. For converse, we can find sequences $x_n \in C_1$ and $y_n \in C_2$ in each group converging to x respectively, but we cannot guarantee that such sequences are the same. Additionally assume that $\mathrm{ri}(C_1) \cap \mathrm{ri}(C_2) \neq \phi$, and for $z \in \mathrm{ri}(C_1) \cap \mathrm{ri}(C_2) = \mathrm{ri}(C_1 \cap C_2)$, $\forall \alpha \in (0, 1)$,

$$\alpha z + (1 - \alpha)x \in \mathrm{ri}(C_1 \cap C_2)$$

by LSP. Let $\alpha_n \searrow 0$ be a sequence in $(0, 1)$, and letting

$$x_n = \alpha_n z + (1 - \alpha_n)x,$$

we get $x_n \in \mathrm{ri}(C_1 \cap C_2) \subseteq C_1 \subseteq C_2$ and $x_n \to x$. □

Example 1.3.16. Again and again, we see a counter-example. Let

$$C_1 = [0, \infty),\ C_2 = (-\infty, 0].$$

Then $\mathrm{ri}(C_1) \cap \mathrm{ri}(C_2) = \phi$, so "=" in (a) does not hold, $\mathrm{ri}(C_1 \cap C_2) = \mathrm{ri}(\{0\}) = \{0\} \neq \mathrm{ri}(C_1) \cap \mathrm{ri}(C_2)$. ("relative") Letting $\tilde{C}_1 = \mathrm{ri}(C_1)$ and $\tilde{C}_2 = \mathrm{ri}(C_2)$, we get

$$\mathrm{cl}(\tilde{C}_1 \cap \tilde{C}_2) = \phi,\ \mathrm{cl}(\tilde{C}_1) \cap \mathrm{cl}(\tilde{C}_2) = \{0\},$$

which becomes a counter-example for (b).

Proposition 1.3.17. *Let $\phi \neq C \subseteq \mathbb{R}^n$ be a nonempty convex set, and $A \in \mathbb{R}^{m\times n}$. If $A^{-1}(\mathrm{ri}(C)) \neq \phi$, then followings hold.*

(a) $\mathrm{ri}\left(A^{-1}(C)\right) = A^{-1}(\mathrm{ri}(C))$.

(b) $\mathrm{cl}\left(A^{-1}(C)\right) = A^{-1}(\mathrm{cl}(C))$.

Sketch of proof. $x \in \mathrm{ri}\left(A^{-1}(C)\right) \Leftrightarrow \forall y \in A^{-1}(C)\ \exists \gamma > 0$ s.t. $x + \gamma(x - y) \in A^{-1}(C) \Leftrightarrow \forall Ay \in C\ \exists \gamma > 0$ s.t. $Ax + \gamma(Ax - Ay) \in C \Leftrightarrow Ax \in \mathrm{ri}(C) \Leftrightarrow x \in A^{-1}(\mathrm{ri}(C))$. Next, $x \in \mathrm{cl}\left(A^{-1}(C)\right) \Rightarrow \exists x_n \in A^{-1}(C)$ s.t. $x_n \to x \Rightarrow \exists Ax_n \in C$ s.t. $Ax_n \to Ax \Rightarrow Ax \in \mathrm{cl}(C)$

implies $\mathrm{cl}\left(A^{-1}(C)\right) \subseteq A^{-1}\left(\mathrm{cl}(C)\right)$. Converse part is clear from

$$\mathrm{ri}\left(A^{-1}(\mathrm{cl}(C))\right) \underset{(*)}{\subseteq} A^{-1}(C) \subseteq A^{-1}(\mathrm{cl}(C)) \subseteq \mathrm{cl}\left(A^{-1}(\mathrm{cl}(C))\right).$$

$(*)$ part comes from

$$\mathrm{ri}\left(A^{-1}\left(\mathrm{cl}(C)\right)\right) = A^{-1}\left(\mathrm{ri}\left(\mathrm{cl}(C)\right)\right) = A^{-1}\left(\mathrm{ri}(C)\right) = A^{-1}\left(\mathrm{ri}(C)\right) \subseteq A^{-1}(C).$$

Proposition 1.3.18 ("Slice" of convex set)**.** *Let $C \subseteq \mathbb{R}^n \times \mathbb{R}^m$ be a convex set. Define*

$$C_x = \{y : (x, y) \in C,\ y \in \mathbb{R}^m\}$$

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

$$D := \{x : C_x \neq \phi\}.$$

Then

$$\mathrm{ri}(C) = \{(x, y) : x \in \mathrm{ri}(D),\ y \in \mathrm{ri}(C_x)\}.$$

In here, C_x is the "slice set" of C at x.

1.3.2 Continuity of convex functions

Proposition 1.3.19. *If $f : \mathbb{R}^n \to (-\infty, \infty]$ is proper convex function, then f is* ***continuous*** *over $\mathrm{ri}(dom(f))$. In particular, if $f : \mathbb{R}^n \to \mathbb{R}$ is convex, (f is clearly prper and therefore) f is continuous (on $\mathbb{R}^n$).*

Proof. WLOG $0 \in \mathrm{ri}(\mathrm{dom}(f))$, and it will be suffice to show that f is continuous at 0. Let $X := \{x : \|x\|_\infty \le 1\} \cap \mathrm{aff}(\mathrm{dom}(f)) \subseteq \mathrm{dom}(f)$ (from $x \in \mathrm{ri}(\mathrm{dom}(f))$, $\exists \delta > 0$ s.t. $\{x : \|x\|_\infty \le \delta\} \cap \mathrm{aff}(\mathrm{dom}(f)) \subseteq \mathrm{dom}(f)$, but WLOG $\delta = 1$). Also define

$$e_i = (\pm 1, \pm 1, \cdots, \pm 1)^\top \in \mathbb{R}^n,\ i = 1, 2, \cdots, 2^n.$$

Then $\|e_i\|_\infty = 1$, and convex hull of $\{e_1, e_2, \cdots, e_{2^n}\}$ is $\{x : \|x\|_\infty \le 1\}$. Thus, all $x \in X$ can be represented as

$$x = \sum_{i=1}^{2^n} \alpha_i e_i,\ \alpha_i \ge 0,\ i = 1, 2, \cdots, 2^n,\ \sum_{i=1}^{2^n} \alpha_i = 1,$$

and hence

$$f(x) = f\left(\sum_{i=1}^{2^n} \alpha_i e_i\right) \le \sum_{i=1}^{2^n} \alpha_i f(e_i) \le \max_{1\le i\le 2^n} f(e_i).$$

Now, for any sequence $\{x_k\} \subseteq X$ s.t. $x_k \xrightarrow[k\to\infty]{} 0$, $x_k \neq 0$, (Actually, we should show $\forall\{x_k\}$ s.t. $x_k \to 0$, $f(x_k) \to 0$ holds, but if $x_k \to 0$, $\|x_k\|_\infty$ becomes small for sufficiently large k, so it's OK only to think in X.) define

$$y_k = \frac{x_k}{\|x_k\|_\infty},\ z_k = -\frac{x_k}{\|x_k\|_\infty}.$$

Then $\|y_k\|_\infty = \|z_k\|_\infty = 1$. Since $\|x_k\|_\infty \le 1$, by convexity,

$$f(x_k) \le (1 - \|x_k\|_\infty) f(0) + \|x_k\|_\infty f(y_k)$$

holds. From $y_k \in X$, $f(y_k)$ is bounded ($\because f(y_k) \le \max f(e_i)$), so $\|x_k\|_\infty f(y_k) \xrightarrow[k\to\infty]{} 0$, which implies

$$\limsup_{k\to\infty} f(x_k) \le \limsup_{k\to\infty} (1 - \|x_k\|_\infty) f(0) + 0 = f(0).$$

However, by convexity again, we get

$$f(0) \le \frac{\|x_k\|_\infty}{1 + \|x_k\|_\infty} f(z_k) + \frac{1}{1 + \|x_k\|_\infty} f(x_k),$$

which implies

$$f(0) \le \liminf_{k\to\infty} \left(\frac{\|x_k\|_\infty}{1 + \|x_k\|_\infty} f(z_k) + \frac{1}{1 + \|x_k\|_\infty} f(x_k)\right) = \liminf_{k\to\infty} f(x_k),$$

via similar logic. Therefore, we get $\lim_{k\to\infty} f(x_k) = f(0)$, which yields continuity. $\square$

Remark 1.3.20. Previous proposition implies that, *real-valued convex function on $\mathbb{R}^n$ is (proper and so) continuous, and hence closed.* Thus remained hard case is extended real-valued function, or function whose domain is not a whole space. For example, consider

$$f(x) = \begin{cases} -x & x > 0 \\ 1 & x = 0 \\ \infty & x < 0 \end{cases}.$$

Then f is proper, convex, but not continuous on $\mathbb{R}$. However, as previous proposition says, it is continuous on $\mathrm{ri}(\mathrm{dom}(f)) = (0, \infty)$.

Proposition 1.3.21. *Let C be a closed interval in $\mathbb{R}$, and $f : C \to \mathbb{R}$ be a closed convex function. Then f is continuous over C.*

Proof. By previous proposition, f is continuous over $\mathrm{ri}(C)$, so remain part is:

Claim. f is continuous at $\bar{x} \in \mathrm{bd}(C) \cap C$. (Since we see on $\mathbb{R}$, rb = bd.)

We can construct $\{x_k\} \subseteq C$ s.t. $x_k \to \bar{x}$, as $x_k = \alpha_k x_0 + (1-\alpha_k)\bar{x}$, for some $x_0 \in \mathrm{ri}(C)$ and $\{\alpha_k\}$ s.t. $\alpha_k \xrightarrow[k\to\infty]{} 0$, $\alpha_k \ge 0$ ($x_k \in C$ is guaranteed from LSP). Then by convexity we get

$$f(x_k) \le \alpha_k f(x_0) + (1-\alpha_k) f(\bar{x}),$$

and hence,

$$\limsup_{k\to\infty} f(x_k) \le f(\bar{x}).$$

Now define an extension $\tilde{f}$ of f as

$$\tilde{f}(x) = \begin{cases} f(x) & x \in C \\ +\infty & o.w. \end{cases}$$

(Why we extended f? Closed function "with whole domain" is lower semi-continuous!) Then $\mathrm{epi}(\tilde{f}) = \mathrm{epi}(f)$, so $\tilde{f}$ is closed, and hence $\tilde{f}$ is lower semi-continuous. It means that,

$$\tilde{f}(\bar{x}) \le \liminf_{k\to\infty} \tilde{f}(x_k) \underset{x_k \in C}{=} \liminf_{k\to\infty} f(x_k).$$

However, from $\bar{x} \in C$, we get $\tilde{f}(\bar{x}) = f(\bar{x})$, which yields

$$\limsup_{k\to\infty} f(x_k) \le f(\bar{x}) \le \liminf_{k\to\infty} f(x_k),$$

and it is what we wanted. □

1.3.3 Closure of function

Now we will define a "closure of function." Recall that: function has one-to-one corresponding set, which is so-called "epigraph." Thus, if we can construct a function whose epigraph is a closure of that of origin function, then defining closure of function might be straightforward. Note that, a set $E \subseteq \mathbb{R}^n$ to be an epigraph, then it should satisfy that

$$\forall (\bar{x}, \bar{w}) \in E,\ \{w : (\bar{x}, w) \in E\} = [a, \infty) \text{ or } \mathbb{R} \text{ for some } a \in \mathbb{R}.$$

Now, for such E, define

$$D := \{x : \exists w \in \mathbb{R} \text{ s.t. } (x, w) \in E\},$$

and

$$\begin{aligned} f : & D \to [-\infty, \infty] \\ & x \mapsto \inf\{w : (x, w) \in E\}. \end{aligned}$$

Extend f to whole domain:

$$\tilde{f}(x) = \begin{cases} f(x) & x \in D \\ \infty & x \notin D \end{cases}$$

Then epigraph of $\tilde{f}$ becomes E, i.e., $\tilde{f}$ is a function we wanted.

Definition 1.3.22 (Closure of a function)**.** *(1) Closure* $\text{cl}(f)$ *of a function* f *is defined as*

$$\left(\text{cl}(f)\right)(x) = \inf\{w : (x, w) \in \text{cl}(\text{epi}(f))\},$$

with convention $\inf \phi = \infty$*. By definition,* $\text{epi}(\text{cl}(f)) = \text{cl}(\text{epi}(f))$*, and hence,* $\text{cl}(f)$ *is closed. Further, if* f *is convex, then* $\text{cl}(f)$ *becomes closed convex.*

(2) "Convex closure" $\breve{\text{cl}}(f)$ *is defined as*

$$\breve{\text{cl}}(f) := \text{cl}(F),$$

where

$$F(x) = \inf\{w : (x, w) \in \mathit{conv}(\text{epi}(f))\}.$$

Note that $\breve{\text{cl}}(f)$ *becomes a closed convex function.*

Figure 1.25: Convex closure $\breve{\text{cl}}(f)$.

Actually, taking closure to a function does not change its minimum.

Proposition 1.3.23. *Let $f : X \to [-\infty, \infty]$ be an arbitrary function. Then,*

$$\inf_{x\in X} f(x) = \inf_{x\in X} (\mathrm{cl}(f))(x) = \inf_{x\in\mathbb{R}^n} (\mathrm{cl}(f))(x) = \inf_{x\in\mathbb{R}^n} F(x) = \inf_{x\in\mathbb{R}^n} \left(\breve{\mathrm{cl}}(f)\right)(x),$$

where

$$F(x) = \inf\{w : (x, w) \in conv(\mathrm{epi}(f))\}.$$

Furthermore, any vector that attains the infimum of f over X also attains the infimum of $\mathrm{cl}f$, F, and $\breve{\mathrm{cl}}f$.

Proof. If $\mathrm{epi}(f) = \phi$, there is nothing to prove ($f \equiv \infty$). Now assume $\mathrm{epi}(f) \neq \phi$. Define

$$f^* = \inf_{x\in\mathbb{R}^n} (\mathrm{cl}(f))(x).$$

Let $\{(\bar{x}_k, \bar{w}_k)\} \subseteq \mathrm{cl}(\mathrm{epi}(f))$ be a sequence s.t. $\bar{w}_k \to f^*$. Then by diagonal argument, $\exists\{(x_k, w_k)\} \subseteq \mathrm{epi}(f)$, $x_k \in X$ s.t. $|w_k - \bar{w}_k| \to 0$, i.e., $w_k \to f^*$. From $f(x_k) \le w_k$, we get

$$\limsup_{k\to\infty} f(x_k) \le f^* \le \mathrm{cl}(f)(x) \le f(x) \ \forall x \in X.$$

The last inequality comes from the definition $\mathrm{cl}(f)(x) = \inf\{w : (x, w) \in \mathrm{cl}(\mathrm{epi}(f))\}$; $f(x) \in \{w : (x, w) \in \mathrm{cl}(\mathrm{epi}(f))\}$. Thus, we get

$$f^* \stackrel{\text{def}}{=} \inf_{x\in\mathbb{R}^n} \mathrm{cl}(f)(x) \le \inf_{x\in X} \mathrm{cl}(f)(x) \le \inf_{x\in X} f(x).$$

However, from $\inf_{x\in X} f(x) \le \liminf_k f(x_k)$, we get

$$f^* \stackrel{\text{def}}{=} \inf_{x\in\mathbb{R}^n} \mathrm{cl}(f)(x) \le \inf_{x\in X} \mathrm{cl}(f)(x) \le \inf_{x\in X} f(x) \le \liminf_{k\to\infty} f(x_k) \le \limsup_{k\to\infty} f(x_k) \le f^*,$$

which yields

$$f^* = \inf_{x\in\mathbb{R}^n} \mathrm{cl}(f)(x) = \inf_{x\in X} \mathrm{cl}(f)(x) = \inf_{x\in X} f(x).$$

Similarly, find $\{(\tilde{x}_k, \tilde{w}_k)\} \subseteq \mathrm{conv}(\mathrm{epi}(f))$ s.t. $\tilde{w}_k \to \inf_{x\in\mathbb{R}^n} F(x)$. Then we can find a convex combination

$$(\tilde{x}_k, \tilde{w}_k) = \sum_{i=1}^m \alpha_i (x_{k_i}, w_{k_i}),\ (x_{k_i}, w_{k_i}) \in \mathrm{epi}(f).$$

Now, by $w_{k_i} \geq f(x_{k_i})$, we get

$$\tilde{w}_k \geq \sum_{i=1}^{m} \alpha_i f(x_{k_i}) \geq \sum_{i=1}^{m} \alpha_i \inf_{x\in X} f(x) = \inf_{x\in X} f(x),$$

and letting $k \to \infty$, we get

$$\inf_{x\in\mathbb{R}^n} F(x) \geq \inf_{x\in X} f(x).$$

However, by definition of $F(x) = \inf\{w : (x,w) \in \mathrm{conv}(\mathrm{epi}(f))\}$ $(f(x) \in \{w : (x,w) \in \mathrm{conv}(\mathrm{epi}(f))\})$, we get $F(x) \leq f(x)$, and therefore

$$\inf_{x\in\mathbb{R}^n} F(x) = \inf_{x\in X} f(x).$$

Now let's see about minimizers. If $\exists x^* \in \underset{x\in X}{\arg\min} f(x)$, then from $f(x) \geq \mathrm{cl}(f)(x)$, we get

$$f(x^*) \geq \mathrm{cl}(f)(x^*).$$

But since $f(x^*) = \inf_{x\in X} f(x) = \inf_{x\in\mathbb{R}^n} \mathrm{cl}(f)(x)$, we get

$$\mathrm{cl}(f)(x^*) = f(x^*),$$

i.e., $x^* \in \underset{x\in\mathbb{R}^n}{\arg\min}\ \mathrm{cl}(f)(x)$. In other words, x^* also achieves infimum of $\mathrm{cl}(f)$. Similarly, from $f(x) \geq F(x) \geq \mathrm{cl}(F)(x) = \breve{\mathrm{cl}}(f)(x)$, we get

$$x^* \in \underset{x\in\mathbb{R}^n}{\arg\min} F(x),\ \ x^* \in \underset{x\in\mathbb{R}^n}{\arg\min}\ \breve{\mathrm{cl}}(f)(x).$$

□

Following proposition says that, "$\mathrm{cl}(f)$ ($\breve{\mathrm{cl}}(f)$, resp.) is the greatest closed (closed convex) function majorized by f."

Proposition 1.3.24. *Let $f : \mathbb{R}^n \to [-\infty, \infty]$ be a function. Then,*

(a) If $g : \mathbb{R}^n \to [-\infty, \infty]$ is closed function dominated by f, i.e., $g(x) \leq f(x)$, then

$$g(x) \leq \mathrm{cl}(f)(x)\ \forall x \in \mathbb{R}^n.$$

(b) If $g : \mathbb{R}^n \to [-\infty, \infty]$ is closed convex function dominated by f, i.e., $g(x) \le f(x)$, then

$$g(x) \le \breve{\mathrm{cl}}(f)(x) \ \forall x \in \mathbb{R}^n.$$

Proof. (a) Note that $g \le f$ implies $\mathrm{epi}(f) \subseteq \mathrm{epi}(g)$. Then

$$\mathrm{epi}(\mathrm{cl}(f)) \overset{\mathrm{def}}{=} \mathrm{cl}(\mathrm{epi}(f)) \subseteq \mathrm{cl}(\mathrm{epi}(g)) = \mathrm{epi}(g)$$

holds, which implies $g \le \mathrm{cl}(f)$. The last equality holds from closedness of g.

(b) Let

$$F(x) = \inf\{w : (x, w) \in \mathrm{conv}(\mathrm{epi}(f))\}.$$

Then

$$\mathrm{epi}(\breve{\mathrm{cl}}(f)) \overset{\breve{\mathrm{cl}}(f)=\mathrm{cl}(F)}{=} \mathrm{cl}(\mathrm{epi}(F)) \overset{\mathrm{def}}{=} \mathrm{cl}(\mathrm{conv}(\mathrm{epi}(f)) \subseteq \mathrm{cl}(\mathrm{conv}(\mathrm{epi}(g))) = \mathrm{epi}(g)$$

holds. The last equality holds since g is closed and convex. □

Now we get following result. Especially, (b) is very useful result in here.

Theorem 1.3.25. *Let $f : \mathbb{R}^n \to [-\infty, \infty]$ be a convex function.*

(a) We get following.

$$\mathrm{cl}(dom(f)) = \mathrm{cl}(dom(\mathrm{cl}(f))), \ \mathrm{ri}(dom(f)) = \mathrm{ri}(dom(\mathrm{cl}(f))).$$

Thus we get

$$\mathrm{cl}(f)(x) = f(x) \ \forall x \in \mathrm{ri}(dom(f)),$$

and consequently,

$$\mathrm{cl}(f) \text{ is proper} \Leftrightarrow f \text{ is proper}.$$

(b) For any $x \in \mathrm{ri}(dom(f))$,

$$\mathrm{cl}(f)(y) = \lim_{\alpha \searrow 0} f(y + \alpha(x - y)) \ \forall y \in \mathbb{R}^n.$$

Proof. (a) First, note that by proposition 1.3.18,

$$\text{ri}(\text{epi}(f)) = \{(x, w) : x \in \text{ri}(\text{dom}(f)),\ f(x) < w\}$$

is obtained ($[f(x), \infty)$ is "the slice" of $\text{epi}(f)$ at x). Using this again, we get

$$\text{ri}(\text{epi}(\text{cl}(f))) = \{(x, w) : x \in \text{ri}(\text{dom}(\text{cl}(f))),\ \text{cl}(f)(x) < w\}.$$

However, by definition and basic topology,

$$\text{ri}(\text{epi}(\text{cl}(f))) = \text{ri}(\text{cl}(\text{epi}(f))) = \text{ri}(\text{epi}(f)),$$

and hence

$$\{(x, w) : x \in \text{ri}(\text{dom}(f)),\ f(x) < w\} = \{(x, w) : x \in \text{ri}(\text{dom}(\text{cl}(f))),\ \text{cl}(f)(x) < w\}.$$

It implies that $\text{ri}(\text{dom}(f)) = \text{ri}(\text{dom}(\text{cl}(f)))$. More precisely, letting A be a projection, we get

$$\text{ri}(\text{dom}(f)) = \text{ri}(A\cdot\text{epi}(f)) = A\cdot\text{ri}(\text{epi}(f)) = A\cdot\text{ri}(\text{epi}(\text{cl}(f))) = \text{ri}(A\cdot\text{epi}(\text{cl}(f))) = \text{ri}(\text{dom}(\text{cl}(f))).$$

Taking closure on both sides we get

$$\text{cl}(\text{dom}(f)) = \text{cl}(\text{dom}(\text{cl}(f))).$$

Now, $\text{ri}(\text{dom}(f)) = \text{ri}(\text{dom}(\text{cl}(f)))$ implies

$$\begin{aligned}\{(x, w) : x \in \text{ri}(\text{dom}(f)),\ f(x) < w\} &= \{(x, w) : x \in \text{ri}(\text{dom}(\text{cl}(f))),\ \text{cl}(f)(x) < w\} \\ &= \{(x, w) : x \in \text{ri}(\text{dom}(f)),\ \text{cl}(f)(x) < w\},\end{aligned}$$

so $f(x) = \text{cl}(f)(x)$ if $x \in \text{ri}(\text{dom}(f)))$. Finally, let's show that $\text{cl}(f)$ is proper if and only if f is proper. First, if f is improper, $\exists x$ s.t. $f(x) = -\infty$, or $f(x) \equiv +\infty$. In these cases, $\text{cl}(f)(x) = -\infty$ ($\because\ \text{cl}(f) \le f$), or $\text{cl}(f) \equiv +\infty$ ($\because\ f \equiv \infty \Rightarrow \text{epi}(f) = \phi \Rightarrow \text{epi}(\text{cl}(f)) = \text{cl}(\text{epi}(f)) = \phi$), respectively. Conversely, assume that $\text{cl}(f)$ is improper. Then by the property of *improper closed convex* function,

$$\text{cl}(f)(x) = \begin{cases} = -\infty & \text{if} x \in \text{dom}(\text{cl}(f)) \\ \infty & \text{o.w.} \end{cases},$$

and hence $\mathrm{cl}(f)(x) = -\infty$ if $x \in \mathrm{ri}(\mathrm{dom}(f))$. Now previous part of this theorem implies that $f(x) = \mathrm{cl}(f)(x) = -\infty$ on $\mathrm{ri}(\mathrm{dom}(f))$, i.e., f is improper.

(b) CASE 1) $y \notin \mathrm{dom}(\mathrm{cl}(f))$. In this case, by definition of effective domain, $\mathrm{cl}(f)(y) = \infty$. Note that, since $\mathrm{cl}(f)$ is closed function defined on the whole domain, it is lower semi-continuous, and hence $\forall \{y_k\}$ s.t. $y_k \xrightarrow[k\to\infty]{} y$,

$$\mathrm{cl}(f)(y) = \infty \le \liminf_{k\to\infty} \mathrm{cl}(f)(y_k),$$

i.e.,

$$\mathrm{cl}(f)(y_k) \xrightarrow[k\to\infty]{} \infty.$$

Now $\mathrm{cl}(f) \le f$ implies $f(y_k) \to \infty$, and therefore,

$$\mathrm{cl}(f)(y) = \infty = \lim_{y_k \to y} f(y_k) = \lim_{\alpha \searrow 0} f(y + \alpha(x - y)).$$

CASE 2) $y \in \mathrm{dom}(\mathrm{cl}(f))$. Define $g : [0,1] \to \mathbb{R}$ as

$$g(\alpha) = \mathrm{cl}(f)(y + \alpha(x - y)).$$

Recall that the assumption is $x \in \mathrm{ri}(\mathrm{dom}(f)) = \mathrm{ri}(\mathrm{dom}(\mathrm{cl}(f)))$. Thus by LSP, for any $y \in \mathrm{dom}(\mathrm{cl}(f))$,

$$y + \alpha(x - y) \in \mathrm{ri}(\mathrm{dom}(\mathrm{cl}(f)))\ \forall \alpha \in (0,1].$$

Hence by (a),

$$\mathrm{cl}(f)(y + \alpha(x - y)) = f(y + \alpha(x - y))\ \forall \alpha \in (0,1].$$

If $\mathrm{cl}(f)(y) = -\infty$, then $\mathrm{cl}(f)$ is improper, so $cl(f)(y + \alpha(x - y)) = -\infty$ from $y + \alpha(x - y) \in \mathrm{dom}(\mathrm{cl}(f))$. Thus,

$$f(y + \alpha(x - y)) = \mathrm{cl}(f)(y + \alpha(x - y)) = -\infty\ \forall \alpha \in (0,1],$$

so the assertion holds. Otherwise, i.e., $\mathrm{cl}(f)(y) > -\infty$, then since $\mathrm{cl}(f)$ is proper, so is f by (a). Hence g is real-valued closed convex function, which is continuous, and therefore,

$$\mathrm{cl}(f)(y) = g(0) = \lim_{\alpha \searrow 0} g(\alpha) = \lim_{\alpha \searrow 0} f(y + \alpha(x - y))$$

holds. $\square$

Now we see a basic result of calculus of closure operations.

Proposition 1.3.26 (Linear composition)**.** *Let $f : \mathbb{R}^n \to [-\infty, \infty]$ be a convex function, and $A \in \mathbb{R}^{m\times n}$ be a linear map. Assume $im(A) \cap \mathrm{ri}(dom(f)) \neq \phi$. Then, for $F(x) = f(Ax)$, we get*

$$\mathrm{cl}(F)(x) = \mathrm{cl}(f)(Ax)\ \forall x \in \mathbb{R}^n,$$

i.e.,

$$\mathrm{cl}(f \circ A) = (\mathrm{cl}(f)) \circ A.$$

Proof. Let $z \in \mathrm{im}(A) \cap \mathrm{ri}(\mathrm{dom}(f))$. Then $\exists y$ s.t. $Ay = z$ (i.e., $y \in A^{-1}z$). Then $\mathrm{dom}(F) = A^{-1}(\mathrm{dom}(f))$ yields $\mathrm{ri}(\mathrm{dom}(F)) = \mathrm{ri}(A^{-1}(\mathrm{dom}(f))) = A^{-1}\left(\mathrm{ri}(\mathrm{dom}(f))\right)$. Now from $z \in \mathrm{ri}(\mathrm{dom}(f))$, we get $y \in A^{-1}\left(\mathrm{ri}(\mathrm{dom}(f))\right) = \mathrm{ri}(\mathrm{dom}(F))$. Thus, by previous theorem,

$$\mathrm{cl}(F)(x) = \lim_{\alpha\searrow 0} F(x + \alpha(y - x)) = \lim_{\alpha\searrow 0} f(Ax + \alpha(Ay - Ax)) = \mathrm{cl}(f)(Ax)$$

holds. In here, the last equality is from previous theorem again, with $Ay \in \mathrm{ri}(\mathrm{dom}(f))$. $\square$

Furthermore note that $F = f \circ A$ is convex. This proposition has a corollary, which is more general result.

Corollary 1.3.27. *Let $f_i : \mathbb{R}^n \to [-\infty, \infty]$ be convex functions for $i = 1, 2, \cdots, m$. Assume that*

$$\bigcap_{i=1}^m \mathrm{ri}(dom(f_i)) \neq \phi.$$

Then for

$$F(x) = f_1(x) + \cdots + f_m(x),$$

F is convex, and further,

$$\mathrm{cl}(F)(x) = \mathrm{cl}(f_1)(x) + \cdots + \mathrm{cl}(f_m)(x).$$

Proof. Let A be a linear map defined as

$$Ax = \underbrace{(x, x, \cdots, x)^\top}_{\#m \text{ of } x},$$

and

$$f(x_1, \cdots, x_m) = f_1(x_1) + \cdots + f_m(x_m).$$

Then $F(x) = f(Ax)$, and from $\mathrm{dom}(F) = \cap_{i=1}^m \mathrm{dom}(f_i)$, we get

$$\bigcap_{i=1}^m \mathrm{ri}(\mathrm{dom}(f_i)) = \mathrm{ri}(\mathrm{dom}(F)) = \mathrm{ri}(A^{-1}(\mathrm{dom}(f))) = A^{-1}\left(\mathrm{ri}(\mathrm{dom}(f))\right) \neq \phi.$$

Then $\exists Ay$ s.t. $Ay \in \mathrm{ri}(\mathrm{dom}(f))$, i.e., $\mathrm{im}(A) \cap \mathrm{ri}(\mathrm{dom}(f)) \neq \phi$. Thus, previous proposition implies

$$\mathrm{cl}(F)(x) = \mathrm{cl}(f)(Ax) = \mathrm{cl}(f)(x, x, \cdots, x).$$

Now note that for $y \in \cap_{i=1}^m \mathrm{ri}(\mathrm{dom}(f_i))$, $(y, y, \cdots, y)^\top = Ay \in \mathrm{ri}(\mathrm{dom}(f))$ holds, and hence

$$\begin{aligned}
\mathrm{cl}(F)(x) &= \mathrm{cl}(f)(Ax) \\
&= \lim_{\alpha \searrow 0} f(Ax + \alpha(Ay - Ax)) \\
&= \lim_{\alpha \searrow 0} \left[f_1(x + \alpha(y - x)) + \cdots + f_m(x + \alpha(y - x))\right] \\
&= \lim_{\alpha \searrow 0} f_1(x + \alpha(y - x)) + \cdots + \lim_{\alpha \searrow 0} f_m(x + \alpha(y - x)) \\
&= \mathrm{cl}(f_1)(x) + \cdots + \mathrm{cl}(f_m)(x)
\end{aligned}$$

is obtained, which is the desired result. The last equality is from $y \in \bigcap_{i=1}^m \mathrm{ri}(\mathrm{dom}(f_i))$. □

Remark 1.3.28. Note that $\cap_{i=1}^m \mathrm{ri}(\mathrm{dom}(f_i)) \neq \phi$ condition is necessary. For example, let $C_1 = (0, \infty)$ and $C_2 = (-\infty, 0)$, and consider

$$f_1 = \delta_{C_1} \text{ and } f_2 = \delta_{C_2},$$

where δ denotes indicator. Recall that: in this course, indicator has ∞ values outside of the set! Then since $C_1 \cap C_2 = \phi$,

$$f_1(x) + f_2(x) = \delta_{C_1}(x) + \delta_{C_2}(x) \equiv \infty,$$

and hence

$$\mathrm{cl}(F)(x) \equiv \infty,$$

where $F(x) = f_1(x) + f_2(x)$. However, since $\mathrm{cl}(C_1) \cap \mathrm{cl}(C_2) = \{0\}$,

$$\mathrm{cl}(f_1)(x) + \mathrm{cl}(f_2)(x) = \delta_{\mathrm{cl}(C_1)}(x) + \delta_{\mathrm{cl}(C_2)}(x) = \delta_{\{0\}}(x) \neq \mathrm{cl}(F)(x).$$

1.4 Recession Cones

Definition 1.4.1. *d is called a **direction of recession** if*

$$x + \alpha d \in C \ \forall x \in C, \ \forall \alpha \geq 0.$$

*(See figure 1.26) Then, **recession cone** R_C of C is defined as the collection of direction of recessions, i.e.,*

$$R_C = \{d : x + \alpha d \in C \ \forall x \in C, \ \forall \alpha \geq 0\}.$$

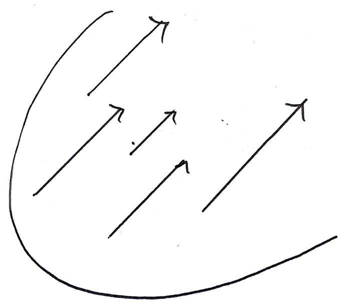


Figure 1.26: Direction of recession

Remark 1.4.2. If d is the direction of recession, then clearly so is λd, if $\lambda > 0$. So clearly R_C is a cone. It says that the term "recession cone" is reasonable. Furthermore, $0 \in R_C$, i.e., recession cone always contains the origin.

Following theorem, so-called "recession cone theorem," shows the behavior of recession cone of closed convex set.

Theorem 1.4.3 (Recession cone theorem)**.** *Let C be a nonempty closed convex set. Then,*

(a) R_C is also closed and convex.

(b) $d \in R_C \Leftrightarrow \exists x \in C$ s.t. $x + \alpha d \in C \ \forall \alpha \geq 0$.

Remark 1.4.4. Note that definition says

$$d \in R_C \Leftrightarrow \forall x \in C \ x + \alpha d \in C \ \forall \alpha \geq 0.$$

This theorem says that if underlying set C is closed and convex, then $d \in R_C$ is guaranteed if we can only find that such point x "exists."

Proof.

Appendix

A Mathematical Background

In this section, we introduce some basic background used oftenly.

A.1 Basic notions

- We often consider *extended real numbers* $\bar{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\}$. Also we define
  - $x \cdot 0 = 0 \ \forall x \in \bar{\mathbb{R}}$
  - $x \cdot \infty = \infty$ if $x > 0$
  - $x \cdot \infty = -\infty$ if $x < 0$
  - $x + \infty = \infty,\ x - \infty = -\infty \ \forall x \in \mathbb{R}$,
  - and we do not allow $\infty - \infty$.
- For nonempty subset X of $\mathbb{R}$, we define $\sup X$ as the smallest $y \in \mathbb{R}$ such that $y \geq x$ for any $x \in X$, and if such y does not exist, we define $\sup X = \infty$. Also, we define $\sup \phi = -\infty$. We can define $\inf X$ similarly.
- If $\sup X := \bar{x}$ is contained in X, we say that $\bar{x} = \max X$. ("maximum is attained")

  If $\inf X := \bar{x}$ is contained in X, we say that $\bar{x} = \min X$. ("minimum is attained")
- Vector space. In this course, we only consider $\mathbb{R}^n$. In here, inner product $\langle x, y \rangle = x^T y$ is defined.
- Also, for $x \in \mathbb{R}^n$, define the notation $x > 0$ or $x \geq 0$ componentwisely. Also define $x > y \Leftrightarrow x - y > 0$.
- Let $f : X \to Y$ be a function. For $U \subseteq X$ and $V \subseteq Y$, we define
  - $f(U) := \{f(x) : x \in U\}$ ("image of U")

- $f^{-1}(V) := \{x \in X : f(x) \in V\}$ ("inverse image of V")

A.2 Linear Algebra

- Let $X, X_1, X_2 \subseteq \mathbb{R}^n$ and λ be a scalar. we define
  - $\lambda X := \{\lambda x : x \in X\}$
  - $X_1 + X_2 := \{x_1 + x_2 : x_1 \in X_1,\ x_2 \in X_2\}$
  - $\overline{x} + X := \{\overline{x}\} + X$ for $\overline{x} \in \mathbb{R}$
  - and $X_1 - X_2 = \{x_1 - x_2 : x_1 \in X_1,\ x_2 \in X_2\}$.
  - To prevent abuse of notation, we will use $X_1 \backslash X_2$ for "set difference."
- If $X_i \subseteq \mathbb{R}^{n_i}$, $i = 1, 2, \cdots, m$, we define "Cartesian Product" as

$$X_1 \times \cdots \times X_m := \{(x_1, \cdots, x_m) : x_i \in X_i,\ i = 1, 2, \cdots, m\} \subseteq \mathbb{R}^{n_1 + \cdots + n_m}.$$

- $S \subseteq \mathbb{R}^n$ is called subspace if $ax + by \in S$ for any $x, y \in S$ and $a, b \in \mathbb{R}$.
- Also, for $\bar{x} \in \mathbb{R}$, $X := \bar{x} + S$ is called an affine set, if S is a subspace.

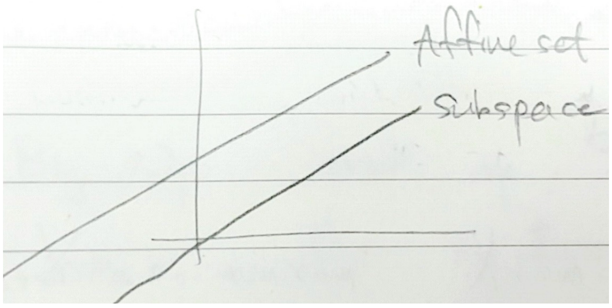


Figure 27: Affine set and subspace

- Facts:
  1. $\exists$ unique subspace associate with an affine set.
  2. $X(\neq \phi)$ is a subspace *if and only if* $0 \in X$ and $\alpha x + (1-\alpha)y \in X$ for any $\alpha \in \mathbb{R}$ and $x, y \in S$.
  3. $X(\neq \phi)$ is an affine set *if and only if* $\alpha x + (1-\alpha)y \in X$ for any $\alpha \in \mathbb{R}$ and $x, y \in S$.
  4. Note that intersection of subspaces is also a subspace.

- span$(x_1, \cdots, x_m)$ is a subspace generated by $x_1, \cdots, x_m$, and it is a set of linear combinations.

- We say that $x_1, \cdots, x_m$ are linearly independent if $\nexists(\alpha_1, \cdots, \alpha_m) \neq 0$ such that $\sum_{k=1}^m \alpha_k x_k = 0$.

- Let S be a nontrivial subspace. Then $\{x_1, \cdots, x_m\}$ is a basis for S if $x_1, \cdots, x_m \in S$, span $(x_1, \cdots, x_m) = S$ and they are linearly independent. In this case, we say $\dim S = m$. Also we define $\dim(\{0\}) = 0$.

- Dimension of the affine set is defined as that of associated subspace. In other words, $\dim(\bar{x} + S) = \dim S$.

- For given a and b, we define $\{x \in \mathbb{R}^n : a^T x = b\}$ as a hyperplane.

- Let $X \subseteq \mathbb{R}^n$. Then $X^\perp := \{y : \langle y, x\rangle = 0 \ \forall x \in X\}$ is a subspace of $\mathbb{R}^n$. In particular, if S is a subspace, then $S^\perp$ is an orthogoanl complement of S. We can say that $\mathbb{R}^n = S \oplus S^\perp$, and $\left(S^\perp\right)^\perp = S$.

- Matrices. Let $A \in \mathbb{R}^{m\times n}$. Then we define $AX := \{Ax : x \in X\}$ and $A^{-1}Y := \{x : Ax \in Y\}$.

- Let $\mathbb{S}^n = \{A \in \mathbb{R}^{n\times n} : A^T = A\}$. Then positive definite matrices are elements of the set

$$\mathbb{S}^n_{++} := \{A \in \mathbb{S}^n : x^T Ax > 0 \ \forall x \in \mathbb{R}^n \backslash \{0\}\},$$

  and denote as $A \succ 0$ if A is s.p.d.. Also, we define a set of nonnegative definite matrices

$$\mathbb{S}^n_{+} := \{A \in \mathbb{S}^n : x^T Ax > 0 \ \forall x \in \mathbb{R}^n \backslash \{0\}\},$$

  and denote as $A \succcurlyeq 0$ if $A \in \mathbb{S}^n_+$.

- If $A \succcurlyeq 0$, then there exists M such that $A = M^T M$.

- For a matrix $A \in \mathbb{R}^{m\times n}$, we define range and null space of A as

$$\mathcal{R}(A) = \{Ax : x \in \mathbb{R}^n\}$$

$$\mathcal{N}(A) = \{x : Ax = 0\}$$

  respectively.

- Rank of matrix A is defined as $rank(A) = \dim(\mathcal{R}(A))$. Note that, $rank(A) = rank(A^T)$, and $\mathcal{R}(A) = \left(\mathcal{N}(A^T)\right)^{\perp}$.

- If $rank(A) = m \wedge n$ we say that A is of full rank.

A.3 Basic Topology

- In here we often use the Euclidean norm $\|x\| = \sqrt{x^T x}$. Then, Cauchy Schwarz inequality $|x^T y| \le \|x\| \cdot \|y\|$ and Pythagoras theorem $\langle x, y\rangle = 0 \Rightarrow \|x+y\|^2 = \|x\|^2 + \|y\|^2$ are known.

- Let $\{x_k\}$ be a sequence in $\mathbb{R}$. We say $\{x_k\}$ converges if $\exists x \in \mathbb{R}$ such that $\forall \epsilon > 0\ \exists K\ s.t.\ \forall k \ge K \Rightarrow |x_k - x| < \epsilon$.

- Also we say that x_k diverges to ∞ if $\forall b \in \mathbb{R}\ \exists K\ s.t.\ \forall k \ge K\ x_k \ge b$.

- $\{x_k\}$ is bounded above if $\exists b$ such that $x_k \le b$ for any k.

- We can define

$$\limsup_{k\to\infty} x_k := \inf_{m\ge1} \sup_{k\ge m} x_k = \lim_{m\to\infty} \sup_{k\ge m} x_k$$

$$\liminf_{k\to\infty} x_k := \sup_{m\ge1} \inf_{k\ge m} x_k = \lim_{m\to\infty} \inf_{k\ge m} x_k.$$

- Note that

$$\inf_{k\ge1} \le \liminf_{k\to\infty} x_k \le \limsup_{k\to\infty} x_k \le \sup_{k\ge1} x_k$$

  holds.

- Also, if for any k $x_k \le y_k$ holds, then $\liminf x_k \le \liminf y_k$ and $\limsup x_k \le \limsup y_k$.

- Moreover,

$$\liminf_{k\to\infty} x_k + \liminf_{k\to\infty} y_k \le \liminf_{k\to\infty}(x_k + y_k)$$

$$\limsup_{k\to\infty} x_k + \limsup_{k\to\infty} y_k \ge \limsup_{k\to\infty}(x_k + y_k)$$

  hold.

- In general, for $\{x_k\} \subseteq \mathbb{R}^n$, we define $x_k rightarrow x$ as $k \to \infty$ if $x_{ki} \to x_i$ as $k \to \infty$. (componentwisely)

- Now we consider a subsequence $\{x_k : k \in \mathcal{K}\}$. x is called limit point if there exists a subsequence such that converges to x.

- Then we get following **Bolzano-Weierstrass Theorem**, *every bounded sequence has at least one limit point.*

- We can define closure $cl(X)$ and interior $int(X)$ of X. Also we can define boundary $bd(X) := cl(X) \backslash int(X)$ of X.

- <u>Facts</u>:

  - The union of a finite collection of closed sets is closed.
  - The intersection of any collection of closed sets is closed.
  - The union of any collection of open sets is open.
  - The intersection of a finite collection of open sets is open.
  - A set is open if and only if all of its elements are interior points.
  - Every subspace of $\mathbb{R}^n$ is closed.
  - A set X is compact if an only if every sequence of elements of X has a subsequence that converges to an element of X.
  - ("Cantor's intersection theorem", or if underlying space is $\mathbb{R}$, "Nested interval theorem") *If $\{X_k\}$ is a sequence of nonempty and compact sets such that $X_{k+1} \subset X_k$ for all k, then the intersection $\bigcap_{k=0}^{\infty} X_k$ is nonempty and compact.*

- Continuity. A function $f : X \to \mathbb{R}^n$ is continuous at x if for any sequence $\{x_k\}$ converges to x, $\lim_k f(x_k) = f(x)$ holds.

- A function $f : X \to \mathbb{R}^n$ is right-continuous (left-continuous) at x if for any sequence $\{x_k\}$ converges to x satisfying $x_k > x$ ($x_k < x$), $\lim_k f(x_k) = f(x)$ holds.

- A real-valued function $f : X \to \mathbb{R}$ is upper semicontinuous (lower semicontinuous) at $x \in X$ if $f(x) \geq \limsup f(x_k)$ ($f(x) \leq \liminf f(x_k)$) for any sequence $\{x_k\}$ in X that converges to x.

- For example, function

$$f(x) = \begin{cases} \sin(1/x) & x \neq 0 \\ 1 & x = 0 \end{cases}$$

  is upper semicontinuous (Figure 28). For more examples, see figure 29.

- <u>Facts</u>:

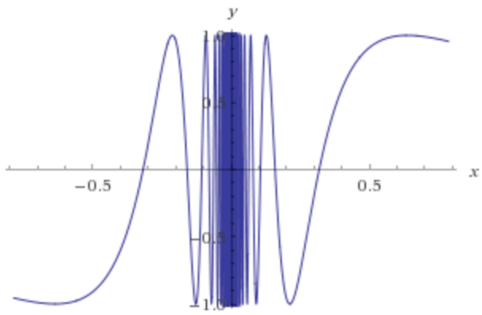


Figure 28: The graph of $y = sin(1/x)$. Image from WolframAlpha.

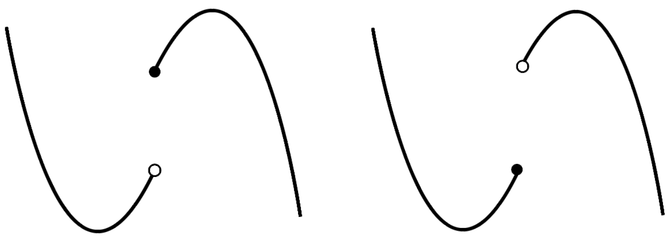


Figure 29: (Left) Such function is upper semicontinuous. (Right) Such function is lower semicontinuous.

- Any vector norm on $\mathbb{R}^n$ is a continuous function.
- Let $f : \mathbb{R}^m \to \mathbb{R}^p$ and $g : \mathbb{R}^n \to \mathbb{R}^m$ be continuous functions. Then $f \circ g$ is also continuous.
- Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be continuous, and Y be an open (closed) subset of $\mathbb{R}^m$. Then $f^{-1}(Y)$ is open (closed).
- Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be continuous, and X be a compact subset of $\mathbb{R}^n$. Then $f(X)$ is compact.

• Following is **Weierstrass' theorem**, or **max-min theorem**: *A continuous function $f : \mathbb{R}^n \to \mathbb{R}$ attains a minimum over any compact subset of $\mathbb{R}^n$.*

Proof. Let $X \subseteq \mathbb{R}^n$ be a compact set. Define a *level set* $V_\gamma = \{x \in X : f(x) \le \gamma\}$, then it is compact since it is bounded and closed. Let $f^* := \inf_{x\in X} f(x) < \infty$. Then for a sequence $\{\gamma_k\}$ such that $\gamma_k \searrow f^*$ and $\gamma_k > f^*$, V_{γ_k} is nonempty, so by Cantor's intersection theorem, $\cap_k V_{\gamma_k}$ is nonempty compact set. Thus, $X^* := \{x \in X : f(x) = f^*\} = \cap_k V_{\gamma_k}$ is nonempty. $\square$