Convex

2022年7月18日

1 Affine sets

$\boldsymbol{x} = (\xi_1, ..., \xi_n) \subset \mathbb{R}^n$
$\langle x, x^* \rangle = \sum_{i=1}^{n} \xi_i \cdot \xi_i^*$
$A : m \times n$ real matrix and linear operator from $\mathbb{R}^n$ to $\mathbb{R}^m$
$f : \mathbb{R}^n \to \mathbb{R}^m : x \mapsto Ax$
$A^* = \overline{A}^\top$: adjoint matrix (随伴行列、エルミート転置行列)
(the transpose matrix and the corresponding adjoint linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$)
$\langle Ax, y^* \rangle = \langle x, A^* y^* \rangle$

Definition 1.1 (affine set)**.** A subset M of $\mathbb{R}^n$ is called *affine set* if

$$\theta x + (1-\theta) y \in M, \qquad \forall x, y \in M, \forall \theta \in \mathbb{R}$$

The empty set $\emptyset$ and the space $\mathbb{R}^n$ itself are affine sets. Also coverd by the definition is the case where M consists of a solitary point.(一点集合も Affine set) An affine set has to contain, along with any two different points, the entire line through those points.

Theorem 1.1. *The subspaces of $\mathbb{R}^n$ are the affine sets which contain the origin.*

Definition 1.2. For $M \subset \mathbb{R}^n$ and $a \in \mathbb{N}$, the translate(平行移動) of M by a is defined to be the set

$$M + a = \{x + a \,|\, x \in M\}$$

An affine set M is said to be parallel to an affine set L if $L = M + a$ for some a.

Theorem 1.2. *Each non-empty affine set M is parallel to a unique subspace L. This L is given by*

$$L = C - C = \{x - y \,|\, x, y \in C\}$$

The dimension of a non-empty affine set is defined as the dimension of the subspace parallel to it. (The dimension of $\emptyset$ is -1 by convention.)

空でないアフィン集合の次元は、それに平行な部分空間の次元と定義される。($\emptyset$ の次元は慣習的に -1)

Affine sets of dimension 0, 1 and 2 are called points, lines andplanes, respectively.

Definition 1.3 (Affine combination)**.** Let $x_1, ..., x_n \in \mathbb{R}^n$, and $\theta_1, ..., \theta_n \in \mathbb{R}(n \in \mathbb{N})$. $\sum_{i=1}^{n} \theta_i x_i$ is an affine combination of $x_1, ..., x_n$ if $\sum_{i=1}^{n} \theta_i = 1$

Theorem 1.3 (Hyperplane)**.** *Given $\beta \in \mathbb{R}$ and a non-zero $\boldsymbol{b} \in \mathbb{R}^n$, the set*

$$H = \{x \in \mathbb{R}^n \mid \langle \boldsymbol{x}, \boldsymbol{b} \rangle = \boldsymbol{\beta}\}$$

is a hyperplane in $\mathbb{R}^n$.

The vector $\boldsymbol{b}$ is called a normal to the hyperplane H.
Every hyper plane has "two sides," loke one's picture of alne in $\mathbb{R}^2$ or a plane in $\mathbb{R}^3$.

Theorem 1.4. *Given $\boldsymbol{b} \in \mathbb{R}^m$ and an $m \times n$ real matrix $\boldsymbol{B}$, the set*

$$M = \{x \in \mathbb{R}^n \mid \boldsymbol{Bx} = \boldsymbol{b}\}$$

is an affine set in $\mathbb{R}^n$. Moreover, every affine set may be represented in this way.

Let $\boldsymbol{b_i}$ is the ith row of $\boldsymbol{B}$, β_i is the i-th component of $\boldsymbol{b}$, and

$$H_i = \{\boldsymbol{x} \mid \langle \boldsymbol{x}, \boldsymbol{b_i} \rangle = \beta_i, \}.$$

Then,

$$M = \{\boldsymbol{x} \mid \langle \boldsymbol{x}, \boldsymbol{b_i} \rangle = \beta_i, i = 1, ..., m\} = \bigcap_{i=1}^{m} H_i.$$

Each H_i is a hyperplane ($\boldsymbol{b_i} \neq \boldsymbol{0}$).

The affine set M in Theorem can be expressed in terms of the vebtors $b'_1, ..., b'_n$ which form the columns of $\boldsymbol{B}$ by

$$M = \{\boldsymbol{x} = (\xi_1, ..., \xi_n) \mid \sum_{i=1}^{n} \xi_i b'_i = \boldsymbol{b}\}$$

任意の集合 $S \in\subset \mathbb{R}^n$ に対し、S を含む（唯一で）最小のアフィン集合が存在する。これをアフィン包 $(affinehull)$ という。

Definition 1.4 (affine hull)**.** A subset C of $\mathbb{R}^n$ is callded *affine hull* if

$$\operatorname{aff} S = \{\sum_{i=1}^{n} \theta_i x_i \mid x_i \in S(i = 1, ..., n), \sum_{i=1}^{n} \theta_i = 1\}.$$

Definition 1.5. A set of $k+1$points $b_0, b_1, \ldots, b_k$ is said *affinely independent* if the set

$$\begin{aligned}\mathrm{aff}\{b_0, b_1, \ldots, b_k\} &= \mathrm{aff}\{b_0 - b_0, b_1 - b_0, \ldots, b_k - b_0\} + b_0 \\ &= \mathrm{aff}\{0, b_1 - b_0, \ldots, b_k - b_0\} + b_0 \\ &= L + b_0 \quad \because L = \mathrm{aff}\{0, b_1 - b_0, \ldots, b_k - b_0\}\end{aligned}$$

is k-dimensional.

定理 1.1 より、L は $b_1 - b_0, \ldots, b_k - b_0$ を含む最小の部分集合と同じになる。L の次元が k になる必要十分条件は $b_1 - b_0, \ldots, b_k - b_0$ が、線形独立になることである。従って、$b_0, b_1, \ldots, b_k$ がアフィン独立である必要十分条件は $b_1 - b_0, \ldots, b_k - b_0$ が線形独立であることである。

よって $x_0, x_1, \ldots, x_k$ がアフィン独立であれば、 $x \in \mathrm{aff}\{b_0, b_1, \ldots, b_k\}$ は次のように表せる。

$$x = \sum_{i=0}^{k} \theta_i x_i, \quad \sum_{i=0}^{k} \theta_i = 1.$$

Definition 1.6 (affine transform)**.** $T : x \to Tx$ from $\mathbb{R}^n$ to $\mathbb{R}^m$ is called *affine transformation* if

$$T((1-\lambda)x + \lambda y) = (1-\lambda)Tx + \lambda Ty, \qquad \forall x, y \in \mathbb{R}^n, \forall \lambda \in \mathbb{R}$$

2 Convex Sets and Cones

Definition 2.1 (convex set)**.** A subset C of $\mathbb{R}^n$ is called *convex* if

$$\theta x + (1-\theta)y \in C, \qquad \forall x, y \in C, \forall \theta \in [0,1]$$

All affine sets are convex.

Definition 2.2 (convex combination)**.** An affine combination with a coefficient $\theta_i \in [0,1]$ iscalled a *convex combination*, i.e.

let $x_1, ..., x_n \in \mathbb{R}^n$, and $\theta_1, ..., \theta_n \in \mathbb{R} (n \in \mathbb{N})$. $\sum_{i=1}^{n} \theta_i \cdot x_i$ is an *combex combination* of $x_1, ..., x_n$ if

$$\sum_{i=1}^{n} \theta_i x_i \quad (\theta_i \in [0,1] (i = 1, ..., n)), \qquad \sum_{i=1}^{n} \theta_i = 1,$$

Definition 2.3 (Half-spaces(半空間))**.** For any non-zero $\boldsymbol{b} \in \mathbb{R}^n$ and any β in $\mathbb{R}$, the sets

$$\{\boldsymbol{x} \mid \langle \boldsymbol{x}, \boldsymbol{b} \rangle \le \beta\}, \qquad \{\boldsymbol{x} \mid \langle \boldsymbol{x}, \boldsymbol{b} \rangle \ge \beta\}$$

are called *closed half-spaces.* The sets

$$\{\boldsymbol{x} \mid \langle \boldsymbol{x}, \boldsymbol{b} \rangle < \beta\}, \qquad \{\boldsymbol{x} \mid \langle \boldsymbol{x}, \boldsymbol{b} \rangle > \beta\}$$

are called *open half-spaces.* All four sets are plainly non-empty and convex.

Theorem 2.1. *The intersection of an arbitrary collection of convex sets is convex.*

A set which can be expressed as the inter section of finitely many closed half spaces of $\mathbb{R}^n$ is called a *polyhedral convex set.* (凸多面体)

$H = \{x \in \mathbb{R}^n \mid \langle \boldsymbol{x}, \boldsymbol{b} \rangle = \boldsymbol{\beta}\}$, $a \in H$ とするとき、半空間 ($\{x \in \mathbb{R}^n \mid \langle \boldsymbol{x}, \boldsymbol{b} \rangle \leq \boldsymbol{\beta}\}$) は

$$C = \{x \in \mathbb{R}^n \mid \langle \boldsymbol{x} - \boldsymbol{a}, \boldsymbol{b} \rangle \leq \boldsymbol{0}\}$$

と表せる。$\boldsymbol{b}$ は点 a における法線ベクトルである。

これは点 a から半空間 ($\{x \in \mathbb{R}^n \mid \langle \boldsymbol{x}, \boldsymbol{b} \rangle \leq \boldsymbol{\beta}\}$) 上の点 a に向かうベクトルが超平面の法線ベクトル b と鈍角をなすことを意味している。

よって、半空間 ($\{x \in \mathbb{R}^n \mid \langle \boldsymbol{x}, \boldsymbol{b} \rangle \leq \boldsymbol{\beta}\}$) は b の反対側に位置する。

一方、半空間 ($\{x \in \mathbb{R}^n \mid \langle \boldsymbol{x}, \boldsymbol{b} \rangle \geq \boldsymbol{\beta}\}$) は b の同じ側に位置する。

Corollary 2.1.1. *Let $b_i \in \mathbb{R}^n$ and $\beta_i \in \mathbb{R}$ for $i \in I$, where I is an arbitrary index set, and consider the set*

$$C = \{x \in \mathbb{R}^n \mid \langle \boldsymbol{x}, \boldsymbol{b_i} \rangle \leq \boldsymbol{\beta_i}, \forall i \in I\}.$$

It is crealy convex.

Theorem 2.2. *A set $C \in \mathbb{R}^n$ is convex if and only if it contains all the convex combinations of its elements.*

($C \in \mathbb{R}^n$ が凸集合 $\Longleftrightarrow$ C の元からなる全ての凸結合を C 自身が含む)

Theorem 2.3 (convex hull)**.** *The convex hull of a given set $X \subset \mathbb{R}^n$ may be defined as the set satisfying any one (and hence all) of the following equivalence conditions.*

1. *The (unique) minimal convex set containing X. (X を含む（唯一の）最小の凸集合)*
2. *The intersection of all convex sets containing X.*

$$convX = \cap\{C : C \text{ is convex set}, X \subset C\}$$

3. *The set of all the convex combinations of points in X. (X に属する点から得られる凸結合全体の成す集合)*

$$convX = \{\sum_{i=1}^{n} \theta_i x_i \in \mathbb{R}^n \mid \exists m \in \mathbb{N},\, \exists x_1, ..., x_m \in X,\, \exists \theta_1, ..., \theta_m \in [0,1],\, \sum_{i=1}^{n} \theta_i = 1\}$$

4. *(Carathéodory's Theorem)The union of all simplices with vertices in X. (X に属する点を頂点とする単体全ての合併)*

$$convX = \{\sum_{i=1}^{n+1} \theta_i x_i \quad | \quad n \in \mathbb{N}, x_i \in X, \theta_i > 0, \sum_{i=1}^{n+1} \theta_i = 1\}$$

A set which is a convex hull of a finite number of points is called a polytope(超多面体).

Definition 2.4. A subset K of $\mathbb{R}^n$ is called *cone* if it is closed under positive scalar multiplication, i.e. $\lambda x \in K$ for $x \in K$ and $\lambda > 0$.

Definition 2.5. A subset K of $\mathbb{R}^n$ is called *convex cone* if it is closed under positive scalar multiplication, i.e. $\lambda x + (1-\lambda)y \in K$ for $x, y \in K$ and $\lambda > 0$.

If K contains the origin, we call it a convex cone.

A convex cone should not necessarily be considered "pointed". A subspace of $\mathbb{R}^n$ is a particularly convex cones. The same is true is the open and closed half-space corresponding to a hyperplane through the origin.

Non-negative *orthant* of $\mathbb{R}^n$ is convex cone.

$$\mathbb{R}^n_+ = \{x = (\xi_1, \ldots, \xi_n) \in \mathbb{R}^n \mid \xi_i \geq 0 \,\text{for}\, i = 1, \ldots, n\} = \{x \in \mathbb{R}^n \mid x \geq 0\}$$
$$\mathbb{R}^n_{>0} = \{x = (\xi_1, \ldots, \xi_n) \in \mathbb{R}^n \mid \xi_i > 0 \,\text{for}\, i = 1, \ldots, n\} = \{x \in \mathbb{R}^n \mid x > 0\}$$

Theorem 2.4. *The intersection of an arbitrary collection of convex cones is a convex cone.*

Corollary 2.4.1. *Let $b_i \in \mathbb{R}^n$ for $i \in I$, where I is an arbitrary index set. Then*

$$K = \{x \in \mathbb{R}^n \mid \langle \boldsymbol{x}, \boldsymbol{b_i} \rangle \leq 0, i \in I\}$$

is a convex cone.

3 Convex Functions

This note is excluded the value $f(x) = -\infty$.

Definition 3.1. A function $f : C \to \mathbb{R}$, where C is a convex set. Then f is *convex* on C if and only if

$$f((1-\lambda)x + \lambda y) \leq (1-\lambda)f(x) + \lambda f(y), \quad \forall x, y \in C, \quad \lambda \in [0,1].$$

Let f : convex function, $f \in C \neq \emptyset$. Extending f to $\mathbb{R}^n$ by setting $f(x) = +\infty$ for $x \notin C \subset \mathbb{R}^n$. Then the above definition is equivalent to the following definition.

Definition 3.2. A function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, not identically $+\infty$, is convex if and only if

$$f((1-\lambda)\boldsymbol{x} + \lambda \boldsymbol{y}) \leq (1-\lambda)f(\boldsymbol{x}) + \lambda f(\boldsymbol{y}), \quad \forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n, \quad \lambda \in [0,1].$$

$f : \mathbb{R}^n \to \mathbb{R}$: *concave* $\Leftrightarrow -f$ is convex.

Definition 3.3. $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}, C \subset \mathbb{R}^n$. We define epigraph of f as

$$\mathrm{epi} f := \{(\boldsymbol{x}, y) \mid \boldsymbol{x} \in C, y \in \mathbb{R}, y \geq f(\boldsymbol{x})\}.$$

Note that $f(x) = +\infty$ for $x \notin C$.

In othe words,
$f:\mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$, not identically equal to $+\infty$. We define epigraph of f is the non-empty set

$$\mathrm{epi} f := \{(\boldsymbol{x}, y) \,|\, \boldsymbol{x} \in \mathbb{R}^n, y \in \mathbb{R}, y \geq f(\boldsymbol{x})\}.$$

Definition 3.4. The *effective domain* of function f on C is the set

$$\mathrm{dom} f = \{\boldsymbol{x} \in \mathbb{R}^n \,|\, \exists y, (\boldsymbol{x}, y) \in \mathrm{epi} f\} = \{\boldsymbol{x} \in \mathbb{R}^n \,|\, f(\boldsymbol{x}) < +\infty\}.$$

Definition 3.5. A convex function f is said to be *proper* if $f(x) < +\infty$ at least one $x \in C$ and $f(x) > -\infty$ for all $x \in C$.
In other words, f is *proper* if its epigraph is non-empty and does not contain "vertical lines".

Let f : convex function, $C \subset \mathbb{R}^n$ is non-empty convex set.
$f \in C$ is proper $\Leftrightarrow$
dom $f = C$ where $f \in C$ is finite.

$f \in \mathbb{R}^n$ is proper $\Leftrightarrow$
$f \in \mathbb{R}^n$ is convex function and $f \in C$ is finite. $\Rightarrow f(x) = \begin{cases} f(x) & (x \in C) \\ +\infty & (x \notin C) \end{cases}$

In this note, we excluded the value $f(x) = -\infty$, so from the definition 3.2, every convex function f is "proper".

Theorem 3.1. $f:\mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$.
f is a convex funciton if and only if epi f *is convex set.*

Theorem 3.2. *Let $f:\mathbb{R}^n \to (-\infty, +\infty], \lambda_i \geq 0$ and $\sum\limits_{i=0}^{k} \lambda_i = 1$. Then f is convex if and only if*

$$f(\sum_{i=0}^{k} \lambda_i x_i) \leq \sum_{i=0}^{k} \lambda_i f(x_i).$$

Neighborhood of x' is defined by

$$B(x', \delta) := \{\bar{x} \,|\, d(x', x) \leq \delta\}.$$

$$V \in \mathcal{N}(x') := \text{the collection of all neiborhoods of } x'.$$

Definition 3.6. (lower limits) Let $f:\mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$ and let x' is a limit point of f. Then the lower limit of function f is defined by

$$\begin{aligned} \liminf_{x\to x'} f(x) &= \lim_{\delta \searrow 0} [\inf_{x\in B(x',\delta)} f(x)] \\ &= \sup_{\delta > 0} [\inf_{x\in B(x',\delta)} f(x)] = \sup_{V\in\mathcal{N}(x')} [\inf_{x\in V} f(x)]. \end{aligned}$$

Definition 3.7. (lower semi-continuous) Let $f:\mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$ and let x' is a limit point of f. Then f is lower semi-continuous at x' if and only if

$$\liminf_{x\to x'} f(x) \geq f(x'), \quad \text{or} \quad \liminf_{x\to x'} f(x) = f(x')$$

Definition 3.8. (level set) The lower level sets $lev_{\leq\alpha}f$ is defined by

$$lev_{\leq\alpha}f := \{x \in \mathbb{R}^n \mid f(x) \leq \alpha\}$$

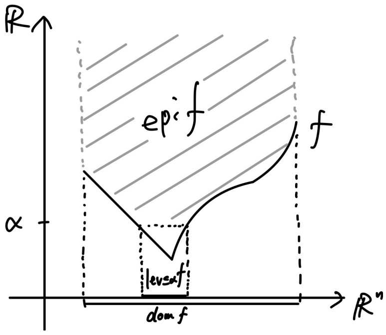

図 1　Level set

Proposition 3.3. *Let arbitrary function* $f : \mathbb{R}^n \to [-\infty, +\infty]$*, the following are equivalent:*

(i) f *is lower semi-continuous at all* $x' \in \mathbb{R}^n$*;*

(ii) $lev_{\leq\alpha}f := \{x^{s\prime} \in \mathbb{R}^n \mid f(x') \leq \alpha\}$ *is closed for every* α*;*

(iii) epif is a closed set in $\mathbb{R}^{n+1}$*.*

Definition 3.9. (level boundedness) A funciton $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is *lower level bounded* if the set $lev_{\leq\alpha}f$ is bounded (possibly empty) for every $\alpha \in \mathbb{R}$.

Theorem 3.4. *(conditions* $\inf f = \min f$*) Suppose* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *is lower semi-continuous, level-bounded and proper. Then the set argminf is nonenpty and compact and the value* $\inf f$ *is finite. In other words,* $\inf f = \min f$*.*

Corollary 3.4.1. *(lower bounds) Suppose* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *is lsc and proper. Then it is bounded from below on each bounded subset of* $\mathbb{R}^n$*. Actually, it is minimum with respect to any compact subset of* $\mathbb{R}^n$ *that satisfies dom f.*

4