02. Convex Sets – 凸集合

By Yang Lin[1] (2023 秋季, @NJU)

Contents

[1]Institute: Nanjing University. Email: linyang@nju.edu.cn.

Preliminaries: linear space

Define set with closed operations ...

Definition: For set V (*e.g.* $\boldsymbol{R}^n$) and field F (*e.g.* $\boldsymbol{R}$), define addition on V

$$\forall x, y \in V \Rightarrow x + y \in V,$$

and scalar multiplication

$$\forall x \in V, \forall \alpha \in F \Rightarrow \alpha x \in V,$$

Also, satisfies (ignored on real numbers)

- $x + y = y + x$
- $1x = x$
- $\alpha(x + y) = \alpha x + \alpha y$
- $(\alpha + \beta)x = \alpha x + \beta x$
- $\alpha(\beta x) = (\alpha\beta)x$

Preliminaries: linear space

Linear subspace: if a subset of linear space $V(F)$ is a linear space, then it is a subspace of $V(F)$

Theorem 1 $W \neq \emptyset$ and $W \subset V$. W is a subspace of $V(F)$ iff

$$\forall x, y \in W \Rightarrow x + y \in W,$$
$$\forall x \in W, \forall \alpha \in F \Rightarrow \alpha x \in W$$

Or equivalently

$$\forall x, y \in W,\ \forall \alpha, \beta \in F \Rightarrow \alpha x + \beta y \in W$$

(only justify it is closed on sums and scalar multiplication)

Theorem 2 $S, T \subset V(F)$ are subspaces, then $S \cap T$ is a subspace, $S \cup T$ is usually not a subspace, $S + T$ is a subspace

$$S + T := \{z | z = x + y, x \in S, y \in T\}$$

Preliminaries: linear space

Theorem 3 a subspace spanned by a set of vectors $x_1, \ldots, x_m$

$$\textbf{Span}\{x_1, \ldots, x_m\} = \left\{ x | x = \sum_{i=1}^{m} \alpha_i x_i, \alpha_i \in F,\ i = 1, \ldots, m \right\}$$

is the minimum subspace containing $x_1, \ldots, x_m$

Element in a linear space

$$x = \sum_{i=1}^{m} \alpha_i x_i, \alpha_i \in F,\ i = 1, \ldots, m$$

Linear space is closed on **linear combination**

Linearly dependent there are α_i some of which are nonzeros, such that $\sum_{i=1}^{m} \alpha_i x_i = 0$

Linearly independent $\sum_{i=1}^{m} \alpha_i x_i = 0$ iff $\alpha_i = 0, \forall i$

Preliminaries: linear space

Dimension of a linear space In space V, there exists $\{x_1, \ldots, x_m\}$ which are linearly independent; and for any set of $m+1$ vectors, the elements are linearly dependent, we call $\{x_1, \ldots, x_m\}$ as the maximal linearly independent system.

$\dim V = m$

Base of a linear space

Important fact on linear dependency and independency:

Vector $y \in \mathbf{Span}\{x_1, \ldots, x_m\}$ has unique coefficients in $y = \sum_{i=1}^m \alpha_i x_i$ iff $x_1, \ldots, x_m$ are linearly independent

系数

Preliminaries: linear space

Theorem subspaces W_1 and W_2, there is

$$\dim(W_1 \cap W_2) \leq \min\{\dim(W_1), \dim(W_2)\}$$
$$\dim(W_1 + W_2) = \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2)$$

Preliminaries: linear map

Definition If V, V' are linear spaces with the same domain F, if $\sigma : V \to V'$ satisfies

$$\sigma(x+y) = \sigma(x) + \sigma(y), \ \forall x, y \in V$$
$$\sigma(\alpha x) = \alpha\sigma(x), \ \forall x \in V, \forall \alpha \in F$$

σ is a linear map from V to V'

A linear map is linear isomorphism map if it is a one-to-one map

A linear map is called linear transformation if $\sigma : V \to V$

Theorem Assume V, V' are subspaces, all linear maps from V to V' consists of set $\mathcal{L}(V, V')$. $\mathcal{L}(V, V')$ is also a subsapces, with the following equations holding

$$(\sigma+\tau)(x) = \sigma(x) + \tau(x), \ \forall x \in V$$
$$(\alpha\sigma)(x) = \alpha\sigma(x), \ \forall x \in V, \forall \alpha \in F$$

Preliminaries: linear map

What is matrix (understand matrix thru linear space)

(1) A set of numbers according to an order (like vector, certainly a linear space)

(2) A set of (row/column) vectors/first-order equations (rank)

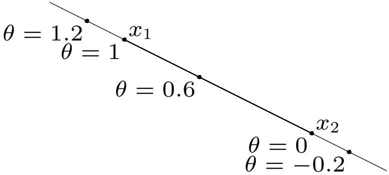
The rank of matrix determines the dimension of the **linear space** that those vectors can span, the solution vector spans the rest (垂直于行向量)

(3) Linear operator/map (linear space)

1 Affine and convex sets

Line Through x_1 and x_2, all points:

$$x = \theta x_1 + (1-\theta)x_2, \ \ (\theta \in \boldsymbol{R})$$



Affine set (仿射集合): contains the line through any two distinct points in the set

- Example (look at analysis in next slide): solution set of linear equations $\{x | Ax = b\}$ (conversely, every affine set can be expressed as solution set of system of linear equations)

1 Affine and convex sets

Alternative explanation of affine set
Assume C is affine, then

$$C = V + x_0 = \{v + x_0 | v \in V\}$$

for some x_0, where V is a subspace
Analysis: For an affine set C and $x_0 \in C$

$$V = C - x_0 = \{x - x_0 | x \in C\}$$

is a subspace (closed under sums and scalar multiplication)

proof ?

1 Affine and convex sets

proof.

Assume $x_0 \in C$ and C is affine, there is:

(1) $0 \in C - x_0$ (do we need that?)

(2) Closed under scalar multiplication

$$x_1 \in C - x_0 \Leftrightarrow x_1 + x_0 \in C$$
$$ax_1 + x_0 = a(x_1 + x_0) + (1 - a)x_0 \in C \Rightarrow ax_1 \in C - x_0$$

(3) Closed under sums

$$x_1 \in C - x_0, x_2 \in C - x_0 \Rightarrow 2x_1 \in C - x_0, 2x_2 \in C - x_0$$
$$x_1 + x_2 + x_0 = \frac{1}{2}(2x_1 + x_0) + \frac{1}{2}(2x_2 + x_0) \in C$$

1 Affine and convex sets

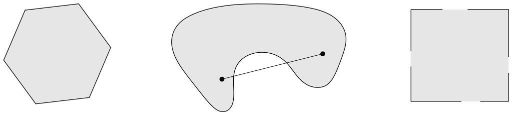
Line segment between x_1 and x_2: all points

$$x = \theta x_1 + (1-\theta)x_2, \ \ (\theta \in [0,1])$$

Convex set: contains line segment between any two points in the set

$$x_1, x_2 \in C, \ \theta \in [0,1] \Rightarrow \theta x_1 + (1-\theta)x_2 \in C$$

Examples: one convex, two non-convex



1 Affine and convex sets

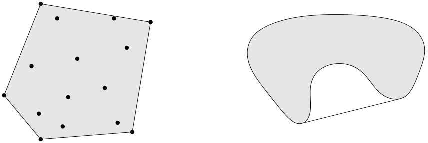
Convex Combination (凸组合) of $x_1, \ldots, x_k$: any point of the form

$$x = \theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_k x_k,$$

with $\theta_1 + \cdots + \theta_k = 1$, and $\theta_i \geq 0$

(Convex set: convex combination lies in the set)

Convex hull (凸包) $\mathbf{conv}\,S$: set of all convex combinations of points in S

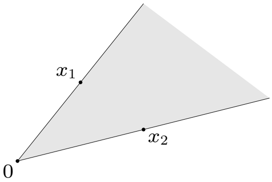


1 Affine and convex sets

Conic (nonnegative) combination of x_1 and x_2: any point of the form

$$x = \theta_1 x_1 + \theta_2 x_2,$$

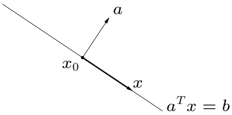
with $\theta_1 \geq 0$, and $\theta_2 \geq 0$



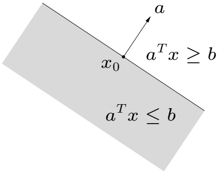
Convex cone: set that contains all conic combinations of points in the set

2 Important examples

Hyperplane: set of the form $\{x | a^T x = b\}$, $a \neq 0$



Halfspace: set of the form $\{x | a^T x \leq b\}$ $(a \neq 0)$



a is the normal vector (确定了法线的方向)

hyperplanes are affine and convex; halfspaces are convex

2 Important examples

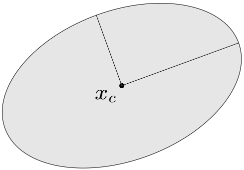
(Euclidean) ball: with center x_c and radius r:

$$B(x_c, r) = \{x \mid \|x - x_c\|_2 \le r\} = \{x_c + ru \mid \|u\|_2 \le 1\}$$

Ellipsoid（椭球）: set of the form

$$\{x \mid \|(x - x_c)^T P^{-1} (x - x_c)\|_2 \le 1\}$$

with $P \in \boldsymbol{S}^n_{++}$ ($i.e.$, P symmetric positive definite)



Other representation: $\{x_c + Au \mid \|u\| \le 1\}$ with A square and nonsingular （非奇异方阵）

2 Important examples

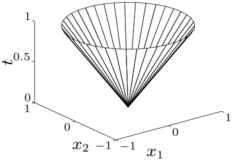
Norm: a function $\|\cdot\|$ that satisfies:

$$\begin{aligned}
&(1) \quad \|x\| \geq 0;\ \ \|x\| = 0 \text{ if and only if } x = 0\\
&(2) \quad \|tx\| = |t|\|x\|, t \in \boldsymbol{R}\\
&(3) \quad \|x + y\| \leq \|x\| + \|y\|
\end{aligned}$$

notation: $\|\cdot\|$ is general (unspecified) norm; $\|\cdot\|_{\texttt{symb}}$ is particular norm

Norm ball(范数球): with center x_c and radius r: $\{x|\ \|x - x_c\| \leq r\}$

Norm cone（范数锥）：$\{(x, t)|\ \|x\| \leq t\}$ (Euclidean norm cone is called second-order cone)



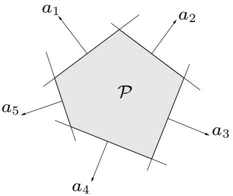
Norm balls and cones are convex

2 Important examples

Polyhedra (多面体): solution set of finitely many linear inequalities and equalities

$$Ax \preceq b, \;\; Cx = d$$

$(A \in R^{m\times n}, C \in R^{p\times n}, \preceq$ is componentwise inequality)



Polyhedron is intersection of finite number of halfspaces and hyperplanes

2 Important examples

Notations:

$\boldsymbol{S}^n$ is set of symmetric $n \times n$ matrices

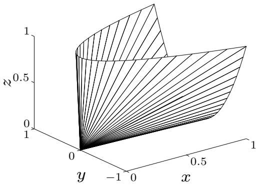
$\boldsymbol{S}^n_+ = \{X \in \boldsymbol{S}^n \mid X \succeq 0\}$: positive semidefinite $n \times n$ matrices

$$X \in \boldsymbol{S}^n_+ \iff z^T X z \geq 0 \text{ for all } z$$

$\boldsymbol{S}^n_+$ is a convex cone

$\boldsymbol{S}^n_{++} = \{X \in \boldsymbol{S}^n \mid X \succ 0\}$: positive definite $n \times n$ matrices

Example: $\begin{bmatrix} x & y \\ y & z \end{bmatrix} \in \boldsymbol{S}^2_+$



2 Important examples

proof.

Assume $A \succeq 0$ and $B \succeq 0$

For any $\theta_1, \theta_2 \geq 0$

$$x^T(\theta_1 A + \theta_2 B)x = \theta_1 x^T A x + \theta_2 x^T B x \geq 0$$

That is $\theta_1 A + \theta_2 B \in \boldsymbol{S}_+^n$

3 Operations that preserve convexity

Practical methods for establishing convexity of a set C:

1. Apply definition

$$x_1, x_2 \in C, 0 \le \theta \le 1 \Rightarrow \theta x_1 + (1-\theta)x_2 \in C$$

2. Show that C is obtained from simple convex sets (hyperplanes, halfspaces, norm balls, ...) by operations that preserve convexity

- Intersection
- Affine functions
- Perspective function – 透视函数
- Linear-fractional functions – 线性分式

3 Operations that preserve convexity

Intersection

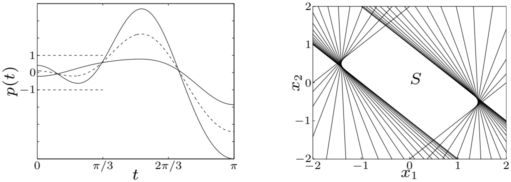
The intersection of (any number of) convex sets is convex

Example

$$S = \{x \in \boldsymbol{R}^m \mid |p(t)| \leq 1 \text{ for } |t| \leq \pi/3\}$$

where $p(t) = x_1 \cos t + x_2 \cos 2t + \cdots + x_m \cos mt$

for $m = 2$



3 Operations that preserve convexity

Affine function

Suppose $f : \boldsymbol{R}^n \to \boldsymbol{R}^m$ is affine ($f(x) = Ax + b$ with $A \in \boldsymbol{R}^{m \times n}$, $b \in \boldsymbol{R}^m$), then:

1. The image of a convex set under f is convex

$$S \subseteq \boldsymbol{R}^n \text{ convex} \Rightarrow f(S) = \{f(x) \mid x \in S\} \text{ convex}$$

2. The inverse image $f^{-1}(C)$ of a convex set under f is convex

$$C \subseteq \boldsymbol{R}^m \text{ convex} \Rightarrow f^{-1}(C) = \{x \in \boldsymbol{R}^n \mid f(x) \in C\} \text{ convex}$$

Examples:

- □ Scaling, translation, projection
- □ Solution set of linear matrix inequality $\{x \mid x_1 A_1 + \cdots + x_m A_m \preceq B\}$ (with $A_i, B \in \boldsymbol{S}^p$)
- □ Hyperbolic cone $\{x \mid x^T P x \le (c^T x)^2, c^T x \ge 0\}$ (with $P \in \boldsymbol{S}^n_+$)

双曲 = 二阶 $\{(z,t) \mid z^T z \le t^2, t \ge 0\}$

3 Operations that preserve convexity

在 $f(x,y)=(\frac{1}{2}x, \sqrt{x})$ 下的像

proof. of Example 2

Define $f(x) := B - A(x) = B - (x_1 A_1 + \cdots + x_m A_m)$

It is affine

Prove set of images $\{f(x) | x \in \mathbf{R}^n, f(x) \succeq 0\}$ is convex

proof. of Example 3

The above set is obtained from the inverse image of $\{(z,t) | z^T z \leq t^2, t \geq 0\}$ through affine function $f(x) = (P^{1/2}x, c^T x)$

3 Operations that preserve convexity

Perspective function $P : \boldsymbol{R}^{n+1} \rightarrow \boldsymbol{R}^n$:

$$P(x,t) = x/t, \ \ \mathbf{dom} P = \{(x,t) | \ t > 0\}$$

Images and inverse images of convex sets under perspective are convex

Linear-fractional function: $f : \boldsymbol{R}^n \rightarrow \boldsymbol{R}^m$:

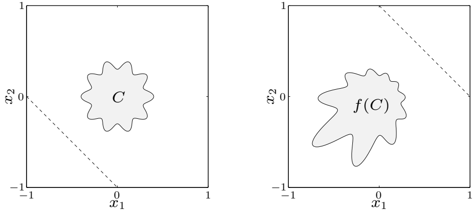
$$f(x) = \frac{Ax+b}{c^T x + d}, \ \ \mathbf{dom}\ f = \{x | \ c^T x + d > 0\}$$

Images and inverse images of convex sets under linear-fractional functions are convex

3 Operations that preserve convexity

Example of a linear-fractional function:

$$f(x) = \frac{1}{x_1 + x_2 + 1} x$$



4 Generalized inequalities – 广义不等式

A convex cone $K \subseteq \boldsymbol{R}^n$ is a proper cone (正常锥) if

- K is closed (contains its boundary)
- K is solid (has nonempty interior)
- K is pointed (contains no line)

Examples:

1. Nonnegative orthant $K = \boldsymbol{R}^n_+ = \{x \in \boldsymbol{R}^n \mid x_i \geq 0, i = 1, \ldots, n\}$
2. Positive semidefinite cone $K = \boldsymbol{S}^n_+$
3. Nonnegative polynomials on $[0, 1]$:

$$K = \{x \in \boldsymbol{R}^n \mid x_1 + x_2 t + x_3 t^2 + \cdots + x_n t^{n-1} \geq 0 \text{ for } t \in [0, 1]\}$$

4 Generalized inequalities – 广义不等式

Generalized inequality defined by a proper cone K:

$$x \preceq_K y \Longleftrightarrow y - x \in K, \quad x \prec_K y \Longleftrightarrow y - x \in \textbf{int } K$$

Examples:

1. Componentwise inequality ($K = \boldsymbol{R}^n_+$)

$$x \preceq_{\boldsymbol{R}^n_+} y \Longleftrightarrow x_i \leq y_i, \quad i = 1, \ldots, n$$

2. Matrix inequality ($K = \boldsymbol{S}^n_+$)

$$X \preceq_{\boldsymbol{S}^n_+} Y \Longleftrightarrow Y - X \text{ positive semidefinite}$$

These two types are so common that we drop the subscript in $\preceq_K$

Many properties of $\preceq_K$ are similar to $\leq$ on $\boldsymbol{R}$, *e.g.*,

$$x \preceq_K y, \ u \preceq_K v \Rightarrow x + u \preceq_K y + v$$

4 Generalized inequalities – 广义不等式

Minimum and minimal elements

$\preceq_K$ is not in general a linear ordering: we can have $x \not\preceq_K y$ and $y \not\preceq_K x$

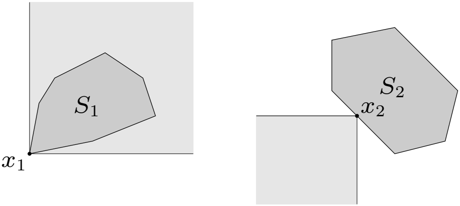
$x \in S$ is the minimum element of S with respect to $\preceq_K$ if

$$y \in S \Rightarrow x \preceq_K y$$

$x \in S$ is the minimal element of S with respect to $\preceq_K$ if

$$y \in S, \ \ y \preceq_K x \Rightarrow y = x$$

Example:

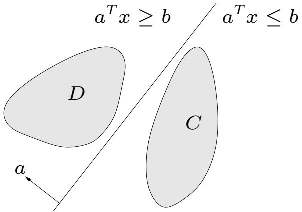


5 Separating and supporting hyperplanes

Separating hyperplane theorem

If C and D are disjoint convex sets, then there exists a $a \neq 0, b$ such that

$$a^T x \leq b \text{ for } x \in C, \;\; a^T x \geq b \text{ for } x \in D$$



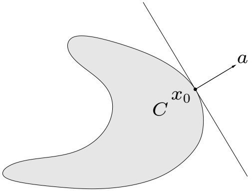
The hyperplane $\{x | a^T x = b\}$ separates C and D

5 Separating and supporting hyperplanes

Supporting hyperplane to set C at boundary point x_0:

$$\{x | a^T x = a^T x_0\}$$

where $a \neq 0$ and $a^T x \leq a^T x_0$ for all $x \in C$



Supporting hyperplane theorem: if C is convex, then there exists a supporting hyperplane at every boundary point of C

6 Dual cones and generalized inequalities

Dual cone of a cone K:

$$K^* = \{y | y^T x \geq 0 \text{ for all } x \in K\}$$

Examples:

- $K = \mathbf{R}^n_+$: $K^* = \mathbf{R}^n_+$
- $K = \mathbf{S}^n_+$: $K^* = \mathbf{S}^n_+$
- $K = \{(x,t) | \|x\|_2 \leq t\}$: $K^* = \{(x,t) | \|x\|_2 \leq t\}$
- $K = \{(x,t) | \|x\|_1 \leq t\}$: $K^* = \{(x,t) | \|x\|_\infty \leq t\}$

First three examples are **self-dual cones**

Dual cones of proper cones are proper, hence define generalized inequalities:

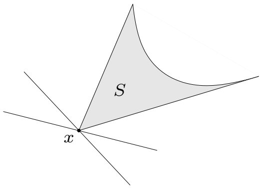
$$y \succeq_{K^*} 0 \Longleftrightarrow y^T x \geq 0 \text{ for all } x \succeq_K 0$$

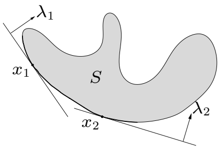
6 Dual cones and generalized inequalities

Minimum and minimal elements via dual inequalities

Minimum element with respect to $\preceq_K$: x is minimum element of S iff for all $\lambda \succ_{K^*} 0$, x is the unique minimizer of $\lambda^T z$ over S

$$K^* = \{y | y^T x \geq 0 \text{ for all } x \in K\}$$





Minimal element with respect to $\preceq_K$: if x minimizes $\lambda^T z$ over S for some $\lambda \succ_{K^*} 0$, then x is minimal

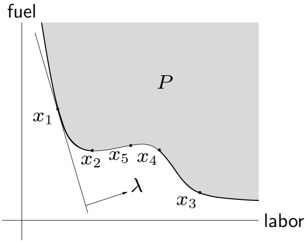
If x is a minimal element of a convex set S, then there exists a nonzero $\lambda \succeq_{K^*} 0$ such that x minimizes $\lambda^T z$ over S

6 Dual cones and generalized inequalities

Optimal production frontier

- Different production methods use different amounts of resources $x \in \boldsymbol{R}^n$
- Production set P: resource vectors x for all possible production methods
- Efficient (Pareto optimal) methods correspond to resource vectors x that are minimal w.r.t. $\boldsymbol{R}^n_+$

example (n =2): x_1, x_2, x_3 are efficient; x_4, x_5 are not



* Homework

《Convex Optimization》Stephen Boyd and Lieven Vandenberghe

2.5, 2.7, 2.10, 2.14, 2.16, 2.20, 2.21, 2.31, 2.39