1 Affine and convex sets

1.1 lines and line segments

line is points of the form $y = \theta x_1 + (1-\theta)x_2$ where $\theta \in R$, and $x_1 \neq x_2$.

line segement corresponds to points where parameter θ is between 0 and 1.

1.2 affine sets

a set is *affine* iff for any $x_1, x_2 \in C$ and $\theta \in R$, we have $\theta x_1 + (1-\theta)x_2 \in C$. This idea can be generalized to more than two points. That is to say, a affine set contains every affine combination of its points: C is an affine set, $x_1, x_2, ..., x_k \in C$, and $\theta_1 + ... + \theta_k = 1$, then the point $\theta_1 x_1 + ... + \theta_k x_k$ also belongs to C.

if C is an affine set and $x_0 \in C$, then the set $V = C - x_0 = \{x \ - \ x_0 \mid x \ \in C\}$ is a subspace. thus any affine set could be expressed as a subspace plus an offset, and the subspace doesnot depend on the choice of x_0.

the set of all affine combination of points in some set C is called the *affine hull* of C:

$$\mathbf{affC} = \{\theta_1 x_1 + ... + \theta_k x_k \mid x_1, x_2, ..., x_k \in C,\ \theta_1 + ... + \theta_k = 1\}. \tag{1}$$

1.3 affine dimension and relative interior

we define the *affine dimension* of a set C as the dimension of its affine hull. if the affine dimension of C is less than n, then the set lies in a affine set $\mathbf{affC} \neq \mathbf{R}^n$. we define *relative interior* if the set C, denoted **relintC**, as its interior relative to **affC**:

$$\mathbf{relintC} = \{x \in C \mid B(x,r) \cap \mathit{affC} \subset C \ for\ some\ r > 0\} \tag{2}$$

where $B(x,r) = \{y | \|y - x\| \leq r\}$. we can then define the *ralative boundary* of a set C as $\mathbf{cl}C$ $\mathbf{relint}C$, where $\mathbf{cl}C$ is the closure of C.

1.4 convex sets

a seet C is *convex* if the line segment between any two points in C lies in C. roughly speaking, a set is convex if every point in the set can be seen by every other point.

we call a point of the form $\theta_1 x_1 + ... + \theta_k x_k$, where $\theta_1 + ... + \theta_k = 1$, and $\theta_i \geq 0$, a *convex combination* of the points $x_1, ..., x_k$. a set is convex iff it contains every convex combination of its points. the *convex hull* of a set C, denoted $\mathbf{conv}C$, is the set of all convex combinations of the points in C. the idea of a convex combination can be generalized to include infinite sums, integrals, and probability distributions. suppose θ_i satisfy

$$\theta_i \geq 0, i = 1, 2, ..., \quad \sum_{i=1}^{\infty} \theta_i = 1, \tag{3}$$

and $x_1, x_2, ... \in C$, where $C \subset \mathbf{R}^n$ is convex, then

$$\sum_{i=1}^{\infty} \theta_i x_i \in C, \tag{4}$$

if the series converges. more generally, suppose $C \subset \mathbf{R}^n$ is convex and x is a random vector with $x \in C$ with probability one, then $\mathbf{E}x \in C$.

1.5 cones

a set is called a *cone*, or *nonnegative homogeneous*, if for every $x \in C$ and $\theta \geq 0$ we have $\theta x \in C$. a point of the form $\theta_1 x_1 + ... + \theta_k x_k$ with $\theta_i \geq 0$ is called *conic combination* or a *nonnegative linear combination* of x_i. a set C is a convex cone iff it contains all conic combinations of its elements. this idea could also be generalized to infinite sums and integrals. the *conic hull* of a set C is the set of all conic combinations of points in C.

2 some important examples

2.1 hyperplanes and halfspaces

a *hyperplane* is a set of the form

$$\{x \mid a^T x = b\}, \tag{5}$$

where $a \in \mathbf{R}^n,\ a \neq 0$, and $b \in \mathbf{R}$. this repersentation can inturn be expressed as

$$\{x \mid a^T(x - x_0) = 0\} = x_0 + a^\perp, \tag{6}$$

where $a^\perp$ denotes the orthogonal complement of a, *i.e.*, the set of all vectors orthogonal to it.

a hyperplane divides $\mathbf{R}^n$ into two *halfspaces*, which a set of the form

$$\{x \mid a^T x \leq b\}, \tag{7}$$

where $a \neq 0$

2.2 euclidean balls and ellipsoids

a *(euclidean) ball* in $\mathbf{R}^n$ has the form

$$B(x_c, r) = \{x \mid \|x - x_c\| \leq r\} = \{x \mid (x - x_c)^T(x - x_c) \leq r^2\}, \tag{8}$$

where $r \geq 0$, and the vector x_c is the *center* of the ball and the saclar r is its *radius*. another common repersentation for the euclidean ball is

$$B(x_c, r) = \{x_c + ru \mid \|u\| \leq 1\}, \tag{9}$$

a euclidean ball is a convex set (use the homogeneity and triangle inequality for $\|\cdot\|_2$)

a related family of convex sets is the *ellipsoids*, which have the form

$$\mathcal{E} = \{x \mid (x - x_c)^T P^{-1}(x - x_c) \leq 1\}, \tag{10}$$

where P is symmetric and positive definite. the vector $x_c \in \mathbf{R}^n$ is the *center* of the ellipsoid. the matrix P determines how far the ellipsoid extends in every direction from x_c by the square root of its eigenvalues. another common repersentation of an ellipsoid is

$$\mathcal{E} = \{x_c + Au \mid \|u\|_2 \leq 1\}, \tag{11}$$

where A is the square and nonsingular, in fact, $A = P^{1/2}$. when the matrix A is symmetric positive semidefinite but not singular, the set is called a *degenerate ellipsoid*, its affine dimension is equal to the rank of A, and it is also convex.

2.3 norm balls and norm cones

a *norm ball* of radius r and center x_c given by $\{x \mid \|x - x_c\| \leq r\}$, is convex. the *norm cone* associated with the norm $\|\cdot\|$ is the set

$$C = \{(x,t) \mid \|x\| \leq t\} \subset \mathbf{R}^{n+1}. \tag{12}$$

it is also a convex set, as the name suggests.

2.4 polyhera

a *polyheron* is defined as the solution set of a finite number of linear equalities and inequalities:

$$\mathcal{P} = \{x \mid a_j^T x \leq b_j,\ j = 1, ..., m,\ c_j^T x = d_j,\ j = 1, ..., p\}. \tag{13}$$

it will be convenient to use the compact notation

$$\mathcal{P} = \{x \mid Ax \preceq b,\ Cx = d\}, \tag{14}$$

where

$$A = \begin{bmatrix} a_1^T \\ \vdots \\ a_m^T \end{bmatrix}, \quad C = \begin{bmatrix} c_1^T \\ \vdots \\ c_p^T \end{bmatrix}, \tag{15}$$

and the symbol $\preceq$ denotes *vector inequality* or *componentwise inequality* in $\mathbf{R}^m$.

2.4.1 simplexes

suppose the $k+1$ points $v+0, ..., v_k \ \in\ \mathbf{R}^n$ are *affinely independent*, then the simplex determined by then is given by

$$C = \mathbf{conv}\{v_0, ..., v_k\} = \{\theta_0 v_0 + \cdots + \theta_k v_k \mid \theta \succeq 0,\ \mathbf{1}^T\theta = 1\}, \tag{16}$$

where $\mathbf{1}$ denotes the vector with all entries one. to show the simplex is a polyheron, we define $y = (\theta_1, ..., \theta_k)$ and

$$B = [v_1 - v_0\ \dots\ v_k - v_0]\ \in\ \mathbf{R}^{nk}, \tag{17}$$

we can say that $x \in C$ iff

$$x = v_0 + By \tag{18}$$

we note that affine independence of the points implies that the matrix B has rank k, therefore there exists a nonsingular matrix $A = (A_1, A_2) \in \mathbf{R}^{nn}$ s.t.

$$AB = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} B = \begin{bmatrix} I \\ 0 \end{bmatrix}. \tag{19}$$

in other words, we have $x \in C$ iff

$$A_2 x = A_2 v_0, \quad A_1 x \succeq A_1 v_0, \quad \mathbf{1}^T A_1 x \leq 1 + \mathbf{1}^T A_1 v_0. \tag{20}$$

2.4.2 convex hull description of polyhedra

the convex hull of a finite set $\{v_1, ..., v_k\}$ is

$$\mathbf{conv}\{v_1, ..., v_k\} = \{\theta_1 v_1 + ..., \theta_k v_k \mid \theta \succeq 0,\ \mathbf{1}^T\theta = 1\}. \tag{21}$$

a generalization of this convex hull description is

$$\{\theta_1 v_1 + ..., \theta_k v_k \mid \theta_1 + \cdots + \theta_m\ = 1,\ \theta_i \geq 0,\ i = 1, ..., k\}, \tag{22}$$

this defines a polyheron, and conversely, every polyheron can be repersented in this form.

2.5 the positive semidefinite cone

we use the notation $\mathbf{S}^n$ $\mathbf{S}^n_+$ $\mathbf{S}^n_{++}$ to be analogous to $\mathbf{R}_+$ $\mathbf{R}_{++}$, then the set $\mathbf{S}^n_+$ is a convex cone.

3 operations that preserve convexity

3.1 intersection

convexity is preserved under intersection: if S_α is convex for every $\alpha \in \mathcal{A}$, then $\cap_{\alpha\in\mathcal{A}} S_\alpha$ is convex.

3.2 affine functions

suppose $S \subset \mathbf{R}^n$ is convex and $f : \mathbf{R}^n \to \mathbf{R}^m$ is an affine function. then the image of S under f,

$$f(S) = \{f(x) \mid x \in S\}, \tag{23}$$

is convex. similarly, the *inverse image* if S under f in convex. two simple examples aer *scaling* and *translation*, *i.e.* αS and $S + \alpha$ are convex. the *projection* of a convex set onto some of its coordinates is convex. the *sum* of two sets $S_1 + S_2$ is convex, if S_1 and S_2 are convex. we can also define the *partial sum* of $S_1, S_2 \in \mathbf{R}^n\mathbf{R}^m$ as

$$S = \{(x, y_1 + y_2) \mid (x, y_1) \in S_1,\ (x, y_2) \in S_2\}, \tag{24}$$

where $x \in \mathbf{R}^n$ and $y_i \in \mathbf{R}^m$. partial sums of convex sets are convex.

3.3 linear-fractional and perspective functions

3.3.1 the perspective function

we define *perspective function* $P : \mathbf{R}^{n+1} \to \mathbf{R}^n$, as $P(z, t) = z/t$, which normalizes vectors so the last component is one, and drop the last component. if $C \subset \mathbf{dom}P$ is convex then its image

$$P(C) = \{P(x) \mid x \in C\} \tag{25}$$

is convex. the imverse image of a convex set under the perspective function is also convex.

3.3.2 linear-fractional functions

a *linear-fractional function* is formed by composing the perspective function with an affine function. suppose $g : \mathbf{R}^n \to \mathbf{R}^{m+1}$ is affine *i.e.*,

$$g(x) = \begin{bmatrix} A \\ c^T \end{bmatrix} x + \begin{bmatrix} b \\ d \end{bmatrix}, \tag{26}$$

the function $f : \mathbf{R}^n \to \mathbf{R}^m$ given by $f = P(g)$, *i.e.*,

$$f(x) = (Ax + b)/(c^T x + d), \quad \mathbf{dom}f = x \mid c^T x + d > 0, \tag{27}$$

is called *linear-fractional* (or *projective*) function. like the perspective function, linear-fractional functions preserve convexity. if we define $mathcal P(z) = \{t(z, 1) \mid t > 0\}$ in $\mathbf{R}^{n+1}$, and

$$Q = \begin{bmatrix} A & b \\ c^T & d \end{bmatrix} \in \mathbf{R}^{(m+1)*(n+1)} \tag{28}$$

then the linear-fractional function can be expressed as

$$f(x) = \mathcal{P}^{-1}(Q\mathcal{P}(x)). \tag{29}$$

4 generalized inequalities

4.1 proper cones and generalized inequalities

a cone $K \subseteq \mathbf{R}^n$ is called a *proper cone* if it satisfies the following:

- K is convex.
- K is closed.
- K is *solid*, which means it has nonempty interior.

- K is *pointed*, which means that it contains no line.

we associate with the proper cone K the partial ordering on $\mathbf{R}^n$ defined by

$$x \preceq_K y \iff y - x \in K. \tag{30}$$

similarly, wedefine an associated strict partial ordering by

$$x \prec_K y \iff y - x \in \mathbf{int}K. \tag{31}$$

when $K = \mathbf{R}_+$, the partial ordering $\preceq_K$ is the usual ordering $\leq$ on $\mathbf{R}$.

4.1.1 properties of generalized inequalities

transitive, reflexive, antisymmetric, preserved under addition, nonnegative scaling, limits.

4.2 minimum and minimal elements

this part is totally the same as *set and graph theory* in PKU.

5 separaing and supporting hyperplanes

5.1 separaing hyperplane theorem

suppose C and D are nonempty disjoint convex sets, then there exists $a \neq 0$ and b such that $a^T x \leq b$ for all $x \in C$ and $a^T x \geq b$ for all $x \in D$.

5.1.1 proof of separaing hyperplane theorem

we assume the (euclidean) *distance* between C and D, defined as

$$\mathbf{dist}(C, D) = inf\{\|u - v\|_2 \mid u \in C,\ v \in D\}, \tag{32}$$

and there exists points $c \in C$ and $d \in D$ that achieve the minimum distance, define $a = d - c, \quad b = \frac{\|d\|_2^2 - \|c\|_2^2}{2}$, we will show that the affine function

$$f(x) = a^T x - b = (d - c)^T (x - (1/2)(c + d)) \tag{33}$$

is nonpositive on C and nonnegative on D, *i.e.*, this hyperplane separates C and D.

5.1.2 strict separation

the separating hyperplane we constructed above satisfies the stronger condition that $a^T x \leq b$ for all $x \in C$ and $a^T x \geq b$ for all $x \in D$, this is called *strict separation.*

5.1.3 converse separaing hyperplane theorem

is not true, that existence of a separaing hyperplane implies that C and D do not intersect, however, if we suppose at least one of these two sets is open, then this is true.

5.2 supporting hyperplanes

suppose $C \in \mathbf{R}^n$, adn x_0 is a point in its boundary $\mathbf{bd}C$, if $a \neq 0$ satisfies $a^T x \leq a^T x_0$ for all $x \in C$, then the hyperplane $\{x \mid a^T x = a^T x_0\}$ is called a *supporting hyperplane* to C at the point x_0. a basic result, called the *supporting hyperplane theorem*, states that for any nonempty convex sets C, and any $x_0 \in \mathbf{bd}C$, there exists a supporting hyperplane to C at x_0.

6 dual cones and generalized inequalities

6.1 dual cones

let K be a cone, the sec

$$K^* = \{y \mid x^T y \geq 0 \ for\ all\ x \in K\} \tag{34}$$

is called the *dual cone* of K. K^* is a cone, and is always convex, even the orginal cone K is not.

dual cone satisfy several properties, such as:

- K^* is closed adn convex
- $K_1 \subseteq K_2$ implies $K_2^* \subseteq K_1^*$
- if K has nonempty interior, then K^* is pointed
- K^{**} is the closure of the convex hull of K

6.2 dual generalized inequalities

suppose that the convex cone K is proper, then it induces a generalized inequality $\preceq_K$, and the *dual* of the generalized inequality $\preceq_{K^*}$

some important properties relating a generalized inequality and its dual are:

- $x \preceq_K y$ iff $\lambda^T x \leq \lambda^T y$ for all $\lambda \succeq_{K^*} 0$
- $x \prec_K y$ iff $\lambda^T x \leq \lambda^T y$ for all $\lambda \succeq_{K^*} 0$, $\lambda \neq 0$.

6.3 minimum adn minimal elements via dual inequalities

6.3.1 dual characterization of minimum element

the *minimum* element: x is the minimum element of S, with respect to the generalized inequality $\preceq_K$, iff for all $\lambda \succ_{K^*} 0$, x is the unique minmizer of $\lambda^T z$ over $z \in S$. this means that for any $\lambda \succ_{K^*} 0$, the hyperplane

$$\{z \mid \lambda^T (z - x) = 0\} \tag{35}$$

is a strict supporting hyperplane to S at x

6.3.2 dual characterization of minimal elements

if $\lambda \succ_{K^*} 0$ and x minimizes $\lambda^T z$ over $z \in S$, then x is minimal, the converse is in general false, except that S is convex.