Math 341: Convex Geometry

Xi Chen

479 Central Academic Building, University of Alberta, Edmonton, Alberta T6G 2G1, CANADA
E-mail address: xichen@math.ualberta.ca
1. Euclidean Geometry

1.1. Vector spaces. Let $\mathbb{R}$ be the set of real numbers. We use $\mathbb{R}^n$ to denote the $n$ dimensional Euclidean space. Every point in $\mathbb{R}^n$ is given by an $n$-tuple $(x_1, x_2, ..., x_n)$, where $x_i \in \mathbb{R}$.

We can define two operations on $\mathbb{R}^n$. One is called vector addition: we define the sum of $(x_1, x_2, ..., x_n) \in \mathbb{R}^n$ and $(y_1, y_2, ..., y_n) \in \mathbb{R}^n$ to be

$$
(1.1) \quad (x_1, x_2, ..., x_n) + (y_1, y_2, ..., y_n) = (x_1 + y_1, x_2 + y_2, x_n + y_n)
$$

The other is called scalar multiplication: we defined the product of $\lambda \in \mathbb{R}$ and $(x_1, x_2, ..., x_n)$ to be

$$
(1.2) \quad \lambda(x_1, x_2, ..., x_n) = (\lambda x_1, \lambda x_2, ..., \lambda x_n)
$$

These two operations obviously have the following properties:

VS1 (Commutativity) For any $u = (x_1, x_2, ..., x_n)$ and $v = (y_1, y_2, ..., y_n) \in \mathbb{R}^n$,

$$
(1.3) \quad u + v = v + u
$$

VS2 (Associativity) For any $u = (x_1, x_2, ..., x_n)$, $v = (y_1, y_2, ..., y_n)$ and $w = (z_1, z_2, ..., z_n) \in \mathbb{R}^n$,

$$
(1.4) \quad (u + v) + w = u + (v + w)
$$

VS3 (Associativity for Scalar Multiplication) For any $\lambda_1, \lambda_2 \in \mathbb{R}$ and $u = (x_1, x_2, ..., x_n) \in \mathbb{R}^n$,

$$
(1.5) \quad (\lambda_1 \lambda_2)u = \lambda_1 (\lambda_2 u)
$$

VS4 (Distribution Law) For any $\lambda \in \mathbb{R}$, $u = (x_1, x_2, ..., x_n)$ and $v = (y_1, y_2, ..., y_n) \in \mathbb{R}^n$,

$$
(1.6) \quad \lambda(u + v) = \lambda u + \lambda v
$$

$\mathbb{R}^n$, equipped with these two operations, is called a vector space over $\mathbb{R}$, or a real vector space. Indeed, any set equipped with two operations that satisfy the axioms VS1-VS4 is a vector space.
1. BASICS

1.1.1. Geometric representation of vectors. A vector \( u = (u_1, u_2, \ldots, u_n) \in \mathbb{R}^n \) can be represented by an “arrow” starting at point \( P = (x_1, x_2, \ldots, x_n) \) and ending at point \( Q = (y_1, y_2, \ldots, y_n) \), where \( y_i = x_i + u_i \) for \( i = 1, 2, \ldots, n \). By convention, we write \( u = PQ \). Let \( O = (0, 0, \ldots, 0) \) be the origin. Since there is a one-to-one correspondence between the vector \( \overrightarrow{OP} \) and the point \( P \), we also call a vector in \( \mathbb{R}^n \) a point from time to time.

Vector addition can be interpreted geometrically by so-called parallelogram criterion: let \( A, B, C, D \) be the vertices (in clockwise order) of a parallelogram. Then \( AB + AD = AC \).

Similarly, here is the geometric interpretation for scalar multiplication: let \( u = PQ \) and \( PR = \lambda u \). Then \( P, Q, R \) are collinear, \( |PR| = |\lambda||PQ| \) and \( Q \) and \( R \) are on the same side of \( P \) iff \( \lambda \geq 0 \).

**Proposition 1.1.** Let \( R \) be the point lying on the line \( PQ \) between \( P \) and \( Q \) and satisfying: \( |PR|/|RQ| = \lambda \). Then

\[
(1.7) \quad \overrightarrow{OR} = t\overrightarrow{OP} + (1-t)\overrightarrow{OQ} \quad \text{with } t = \frac{1}{1+\lambda}
\]

**Proof.** We see from the assumption that \( \overrightarrow{PR} = \lambda \overrightarrow{RQ} \). And since \( \overrightarrow{PR} = \overrightarrow{OR} - \overrightarrow{OP} \) and \( \overrightarrow{RQ} = \overrightarrow{OQ} - \overrightarrow{OR} \), we have \( \overrightarrow{OR} - \overrightarrow{OP} = \lambda (\overrightarrow{OQ} - \overrightarrow{OR}) \). And (1.7) follows. \( \square \)

Let \( V \) be a vector space (over \( \mathbb{R} \)).

1.1.2. Linear dependence and dimension. We say that \( v_1, v_2, \ldots, v_k \in V \) are linearly dependent if there exist \( \lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{R} \), not all zero, such that \( \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_k v_k = 0 \); and \( v_1, v_2, \ldots, v_k \) are linearly independent if they are not linearly dependent.

**Proposition 1.2.** Any \( m > n \) vectors \( v_1, v_2, \ldots, v_m \) in \( \mathbb{R}^n \) are linearly dependent.

**Proof.** Let \( v_i = (a_{i1}, a_{i2}, \ldots, a_{im}) \) for \( i = 1, 2, \ldots, m \) and \( A = (a_{ij})_{m \times n} \). Using Gaussian elimination, \( A \) can be turn into an upper-triangular matrix \( B = (b_{ij})_{m \times n} \) after a series of row operations. That is, there exists a nonsingular matrix \( C = (c_{ij})_{m \times m} \) such that \( CA = B = (b_{ij})_{m \times n} \), where \( b_{ij} = 0 \) for \( i > j \). Therefore,

\[
(1.8) \quad c_{m1}v_1 + c_{m2}v_2 + \cdots + c_{mm}v_m = (b_{m1}, b_{m2}, \ldots, b_{mn}) = 0
\]

Since \( C \) is nonsingular, i.e., \( \det(C) \neq 0 \), \( c_{m1}, c_{m2}, \ldots, c_{mm} \) cannot be all zero. Hence \( v_1, v_2, \ldots, v_m \) are linearly dependent. \( \square \)

We say that \( v \in V \) is a linear combination of \( v_1, v_2, \ldots, v_k \in V \) if there exist \( \lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{R} \) such that \( v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_k v_k \). Let \( S \) be a subset of \( V \). We say that \( S \) generates \( V \) if every vector in \( V \) is a linear combination of some vectors in \( S \). We say that \( S \) is a basis of \( V \) if \( S \) generates \( V \) and \( v_1, v_2, \ldots, v_k \) are linearly independent for any finite subset \( \{v_1, v_2, \ldots, v_k\} \subset S \).
THEOREM 1.3. Let $S_1$ and $S_2$ be two bases of $V$. Then $|S_1| = |S_2|$.

PROOF. We will only prove the theorem when $S_1$ and $S_2$ are finite as that is the only case we need here. Let $S_1 = \{u_1, u_2, ..., u_n\}$ and $S_2 = \{v_1, v_2, ..., v_m\}$. Since $S_1$ generates $V$, $v_j$ is a linear combination of $u_1, u_2, ..., u_n$, i.e., $v_j = a_{1j}u_1 + a_{2j}u_2 + ... + a_{nj}u_n$ for some $a_{ij} \in \mathbb{R}$. Suppose that $m > n$.

By Proposition 1.2, the $m$ vectors $(a_{1j}, a_{2j}, ..., a_{nj})$ are linear dependent. Hence there exist $\lambda_1, \lambda_2, ..., \lambda_m \in \mathbb{R}$, not all zero, such that

$$
\sum_{j=1}^{m} \lambda_j (a_{1j}, a_{2j}, ..., a_{nj}) = 0
$$

By (1.9), we have

$$
\sum_{j=1}^{m} \lambda_j v_j = \sum_{j=1}^{m} \lambda_j (a_{1j}u_1 + a_{2j}u_2 + ... + a_{nj}u_n) = 0
$$

and hence $v_1, v_2, ..., v_m$ are linear dependent. Contradiction. Therefore, $m \leq n$. Similarly, $n \leq m$. Therefore, $m = n$ and $|S_1| = |S_2|$. \hfill \Box

Let $S$ be a basis of $V$ and we define the dimension of vector space $V$ to be $\dim V = |S|$. By the above theorem, $\dim V$ is independent of the choice of basis $S$. To emphasize the fact $V$ is a vector space over $\mathbb{R}$, we write $\dim_\mathbb{R} V$ for the dimension of $V$ over $\mathbb{R}$. Obviously, $\dim_\mathbb{R} \mathbb{R}^n = n$.

1.1.3. Linear subspaces. A subset $W \subset V$ is a (linear) subspace of $V$ if $W$ is closed under vector addition and scalar multiplication, i.e., $u + v \in W$ for any $u, v \in W$ and $\lambda u \in W$ for any $u \in W$ and $\lambda \in \mathbb{R}$. A subspace $W \subset V$ is itself a vector space (over $\mathbb{R}$).

PROPOSITION 1.4. $W \subset V$ is a linear subspace if and only if $u + \lambda v \in W$ for any $u, v \in W$ and $\lambda \in \mathbb{R}$.

PROOF. “$\Rightarrow$”: Since $W$ is a subspace, $\lambda v \in W$ and hence $u + \lambda v \in W$.

“$\Leftarrow$”: Since $u + \lambda v \in W$ for all $u, v \in W$ and $\lambda \in \mathbb{R}$, $u + v \in W$ by letting $\lambda = 1$ and $\lambda v \in W$ by letting $u = 0$. \hfill \Box

For two subsets $A, B \subset V$ and $\lambda \in \mathbb{R}$, we define

$$
A + B = \{a + b : a \in A, b \in B\} \text{ and } \lambda A = \{\lambda a : a \in A\}
$$

If $A = \{x\}$ consists of a single vector, we often write $x + B$ for $A + B$ and we call $x + B$ a translate of $B$. The following are obvious.

PROPOSITION 1.5. Let $A, B, C \subset V$ and $\lambda, \mu \in \mathbb{R}$. Then

1. $A + B = B + A$;
2. $(A + B) + C = A + (B + C)$;
3. $\lambda (A + B) = \lambda A + \lambda B$;
4. $(\lambda + \mu)A = \lambda A + \mu A$;
5. $(\lambda \mu)A = \lambda (\mu A)$;
6. $A$ is a subspace if and only if $A + A = A$ and $\lambda A = A$ for any $\lambda \neq 0$. 

(7) If $A$ and $B$ are subspaces, so is $A + B$.
(8) The intersection of any collection of subspaces is a subspace.

Let $V$ and $W$ be two vector spaces. We call a map $f : V \to W$ a linear map or transformation if $f(x + y) = f(x) + f(y)$ and $f(\lambda x) = \lambda f(x)$ for all $x, y \in V$ and $\lambda \in \mathbb{R}$.

1.1.4. Affine dependence and affine subspaces. The translate of a linear subspace of $V$ is called an affine subspace. Let $W \subset V$ be a linear subspace and $x + W$ be an affine subspace. The dimension of $x + W$ is defined to be the dimension of $W$. An affine subspace of dimension 1 is called a line. An affine subspace in $V$ of dimension $\dim V - 1$ is called a hyperplane. An affine subspace of dimension $k$ is called a $k$-plane.

**Proposition 1.6.**

(1) Let $S \subset V$ be an affine subspace $V$ and let $y \in S$. Then $S - y = (-y) + S$ is a linear subspace of $V$.

(2) The intersection of any collection of affine subspaces is an affine subspace.

**Proof.** For (1), let $S = x + W$ for some $x \in V$ and linear subspace $W \subset V$. Then $y = x + w$ for some $w \in W$. Then $S - y = x + W - (x + w) = W - w$. We claim that $W - w = W$. For any $u \in W$, $u + w \in W$ and hence $u = (u + w) - w \in W - w$; therefore $W \subset W - w$. And since $u - w \in W$, $W - w \subset W$. Therefore, $S - y = W - w = W$ is a linear subspace.

For (2), let $\{S_i \subset V\}_{i \in I}$ be a collection of affine subspaces of $V$. If $\cap_{i \in I} S_i$ is empty, we are done. If not, let $y \in \cap_{i \in I} S_i$. By (1), $W_i = S_i - y$ is a linear subspace. We claim that

\[ \bigcap_{i \in I} S_i = \bigcap_{i \in I} (y + W_i) = y + \bigcap_{i \in I} W_i \]

and then $\cap_{i \in I} S_i$ is an affine subspace since $\cap_{i \in I} W_i$ is a linear subspace.

To prove (1.12), let $x \in \cap_{i \in I} S_i$. Then $x \in S_i = y + W_i$. Hence $x - y \in W_i$ for all $i \in I$. Consequently,

\[ x - y \in \bigcap_{i \in I} W_i \Rightarrow x \in y + \bigcap_{i \in I} W_i \Rightarrow \bigcap_{i \in I} S_i \subset y + \bigcap_{i \in I} W_i \]

Let $w \in \cap_{i \in I} W_i$. Then $y + w \in S_i$ for all $i \in I$. Therefore, $y + w \in \cap_{i \in I} S_i$ and $y + \cap_{i \in I} W_i \subset \cap_{i \in I} S_i$. And (1.12) follows. $\square$

We say vectors $u_1, u_2, \ldots, u_m \in V$ are affinely dependent if there exist $\lambda_1, \lambda_2, \ldots, \lambda_m \in \mathbb{R}$, not all zero, such that $\lambda_1 + \lambda_2 + \ldots + \lambda_m = 0$ and $\lambda_1 u_1 + \lambda_2 u_2 + \ldots + \lambda_m u_m = 0$. We say $u$ is an affine combination of $u_1, u_2, \ldots, u_m$ if there exist $\lambda_1, \lambda_2, \ldots, \lambda_m \in \mathbb{R}$ such that $\lambda_1 + \lambda_2 + \ldots + \lambda_m = 1$ and $u = \lambda_1 u_1 + \lambda_2 u_2 + \ldots + \lambda_m u_m$.

**Proposition 1.7.** Let $v_1, v_2, \ldots, v_m \in \mathbb{R}^n$ with $v_j = (a_{1j}, a_{2j}, \ldots, a_{nj})$ for $j = 1, 2, \ldots, m$ and let $u_j = (a_{1j}, a_{2j}, \ldots, a_{nj}) \in \mathbb{R}^{n+1}$.

(1) $v_1, v_2, \ldots, v_m$ are affinely dependent if and only if $u_1, u_2, \ldots, u_m$ are linearly dependent.
(2) Let \( v = (a_1, a_2, ..., a_n) \) and \( u = (1, a_1, a_2, ..., a_n) \). Then \( v \) is an affine combination of \( v_1, v_2, ..., v_m \) if and only if \( u \) is a linear combination of \( u_1, u_2, ..., u_m \).

**Proof.** For (1),

\[
\begin{align*}
\text{\( u_1, u_2, ..., u_m \) linearly dependent} & \iff \exists \lambda_1, \lambda_2, ..., \lambda_m \in \mathbb{R}, \text{ not all zero, such that} \\
0 &= \sum_{j=1}^{m} \lambda_j u_j = \sum_{j=1}^{m} \lambda_j (1, a_{1j}, a_{2j}, ..., a_{nj}) \\
&\iff \sum_{j=1}^{m} \lambda_j = 0 \text{ and } \sum_{j=1}^{m} \lambda_j v_j = \sum_{j=1}^{m} \lambda_j (a_{1j}, a_{2j}, ..., a_{nj}) = 0 \\
&\iff v_1, v_2, ..., v_m \text{ affinely dependent}
\end{align*}
\]

(1.14)

For (2),

\[
\begin{align*}
\text{\( u \) is a linear combination of \( u_1, u_2, ..., u_m \)} & \iff \exists \lambda_1, \lambda_2, ..., \lambda_m \in \mathbb{R} \text{ such that} \\
(1, a_1, a_2, ..., a_n) = u &= \sum_{j=1}^{m} \lambda_j u_j = \sum_{j=1}^{m} \lambda_j (1, a_{1j}, a_{2j}, ..., a_{nj}) \\
&\iff \sum_{j=1}^{m} \lambda_j = 1 \text{ and } \sum_{j=1}^{m} \lambda_j v_j = \sum_{j=1}^{m} \lambda_j (a_{1j}, a_{2j}, ..., a_{nj}) = (a_1, a_2, ..., a_n) = v \\
&\iff v \text{ is an affine combination of } v_1, v_2, ..., v_m
\end{align*}
\]

(1.15)

\[\square\]

**Corollary 1.8.** Any \( m > n + 1 \) vectors in \( \mathbb{R}^n \) are affinely dependent.

A set \( S \) is said to be affine if every affine combination of two points in \( S \) belongs to \( S \), i.e., \( \lambda x + (1 - \lambda)y \in S \) for any \( x, y \in S \).

**Proposition 1.9.** A set \( S \) is affine if and only if every affine combination of points of \( S \) lies in \( S \).

**Proof.** “\( \Leftarrow \)” is trivial.

“\( \Rightarrow \)” : We prove by induction on \( n \) that \( \lambda_1 v_1 + \lambda_2 v_2 + ... + \lambda_n v_n \in S \) for any \( v_1, v_2, ..., v_n \in S \) and \( \lambda_1 + \lambda_2 + ... + \lambda_n = 1 \). This holds for \( n = 2 \) by definition. Suppose that it holds for \( n < k \). When \( n = k > 2 \). At least one of \( \lambda_1, \lambda_2, ..., \lambda_k \) is not equal to 1. Without the loss of generality, assume that...
1. BASICS

\( \lambda_k \neq 1. \) Then \( \lambda_1 + \lambda_2 + ... + \lambda_{k-1} \neq 0. \) Therefore,

\[
\lambda_1 v_1 + \lambda_2 v_2 + ... + \lambda_{k-1} v_{k-1} + \lambda_k v_k
\]

(1.16)

\[
= (\lambda_1 + \lambda_2 + ... + \lambda_{k-1}) \left( \frac{\lambda_1}{\lambda_1 + \lambda_2 + ... + \lambda_{k-1}} v_1 + \frac{\lambda_2}{\lambda_1 + \lambda_2 + ... + \lambda_{k-1}} v_2 + ... + \frac{\lambda_{k-1}}{\lambda_1 + \lambda_2 + ... + \lambda_{k-1}} v_{k-1} \right) + \lambda_k v_k
\]

By induction hypothesis,

\[
v = \frac{\lambda_1}{\lambda_1 + \lambda_2 + ... + \lambda_{k-1}} v_1 + \frac{\lambda_2}{\lambda_1 + \lambda_2 + ... + \lambda_{k-1}} v_2 + ... + \frac{\lambda_{k-1}}{\lambda_1 + \lambda_2 + ... + \lambda_{k-1}} v_{k-1} \in S
\]

(1.17)

Therefore, \( \lambda_1 v_1 + \lambda_2 v_2 + ... + \lambda_k v_k = (\lambda_1 + \lambda_2 + ... + \lambda_{k-1})v + \lambda_k v_k \in S. \)

**Theorem 1.10.** \( S \subseteq V \) is affine if and only if \( S \) is an affine subspace.

**Proof.** “\( \Rightarrow \)” If \( S \) is empty, there is nothing to prove. Suppose that \( v \in S. \) Let \( W = S - v. \) We will show that \( W \) is a linear subspace of \( V \) and then it follows that \( S \) is an affine subspace.

Let \( x, y \in S \) and \( \lambda \in \mathbb{R}. \) Then

\[
(x - v) + \lambda(y - v) = (x + \lambda y - \lambda v) - v
\]

(1.18)

Note that \( x + \lambda y - \lambda v \) is an affine combination of \( x, y \) and \( v. \) Therefore, \( x + \lambda y - \lambda v \in S, \) \((x - v) + \lambda(y - v) \in W \) and \( W \) is a linear subspace.

“\( \Leftarrow \)” Suppose that \( S \) is an affine subspace. Let \( S = v + W, \) where \( W \) is a linear subspace of \( V. \) For \( x, y \in W \) and \( \lambda \in \mathbb{R}, \)

\[
\lambda(v + x) + (1 - \lambda)(v + y) = v + (\lambda x + (1 - \lambda)y) \in v + W = S
\]

(1.19)

Therefore, \( S \) is affine.

By the above theorem, affine subspaces and affine sets are the same things. Therefore, we will use these two terms interchangeably from now on.

The affine hull of a set \( S \) is the intersection of all the affine sets which contain \( S, \) denoted by \( \text{aff}(S). \) Alternatively, the affine hull of \( S \) can be defined as the smallest affine set that contains \( S \) in the sense that \( \text{aff}(S) \subseteq T \) for any affine set \( T \supset S. \)

**Proposition 1.11.** The affine hull of \( S \) consists precisely of all affine combinations of points of \( S. \) That is,

\[
\text{aff}(S) = T = \{ \lambda_1 x_1 + \lambda_2 x_2 + ... + \lambda_m x_m : \lambda_1 + \lambda_2 + ... + \lambda_m = 1, \ x_1, x_2, ..., x_m \in S \}
\]

(1.20)

**Proof.** By Proposition 1.9, \( \lambda_1 x_1 + \lambda_2 x_2 + ... + \lambda_m x_m \in \text{aff}(S) \) for all \( \lambda_1 + \lambda_2 + ... + \lambda_m = 1, \ x_1, x_2, ..., x_m \in S. \) Therefore, \( T \subseteq \text{aff}(S). \)
On the other hand, for any \( \lambda \in \mathbb{R} \), \( x = \alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_m x_m \) and \( y = \beta_1 y_1 + \beta_2 y_2 + \ldots + \beta_n y_n \in T \) with \( \alpha_1 + \alpha_2 + \ldots + \alpha_m = \beta_1 + \beta_2 + \ldots + \beta_n = 1 \) and \( x_1, x_2, \ldots, x_m, y_1, y_2, \ldots, y_n \in S \),

\[
\lambda x + (1 - \lambda) y = \sum_{i=1}^{m} \lambda \alpha_i x_i + \sum_{j=1}^{n} (1 - \lambda) \beta_j y_j
\]

Since

\[
\sum_{i=1}^{m} \lambda \alpha_i + \sum_{j=1}^{n} (1 - \lambda) \beta_j = 1
\]

\( \lambda x + (1 - \lambda) y \) is an affine combination of \( x_1, x_2, \ldots, x_m, y_1, y_2, \ldots, y_n \). Therefore, \( \lambda x + (1 - \lambda) y \in T \) and \( T \) is affine. It follows from the definition of \( \text{aff}(S) \) that \( \text{aff}(S) \subset T \). Hence \( T = \text{aff}(S) \).

Using the concept of affine hull, we have a (rudimentary) notion of dimensions of subsets in a vector space. The dimension of a set \( S \subset V \) is the dimension of its affine hull \( \text{aff}(S) \). Note that this is not a widely-accepted notion of dimension. For example, take \( S = \{p, q\} \) consisting of two distinct points \( p \) and \( q \). Then \( \text{aff}(S) \) is the line joining \( p \) and \( q \). Then \( \dim(S) = \dim(\text{aff}(S)) = 1 \). However, our intuition tells us \( S \) should have dimension 0; a better notion of dimension should give \( \dim(S) = 0 \). However, this is the definition we will use throughout this notes.

1.2. Topology on \( \mathbb{R}^n \). The inner product \( \langle u, v \rangle \) of \( u = (x_1, x_2, \ldots, x_n) \) and \( v = (y_1, y_2, \ldots, y_n) \in \mathbb{R}^n \) is defined to be

\[
\langle u, v \rangle = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n
\]

Alternatively, we write \( \langle u, v \rangle = u \cdot v \). It is easy to check the following.

**Proposition 1.12.** Let \( u, v, w \in \mathbb{R}^n \) and \( \lambda \in \mathbb{R} \).

1. \( \langle u, u \rangle \geq 0 \) and the equality holds if and only if \( u = 0 \);
2. \( \langle u, v \rangle = \langle v, u \rangle \);
3. \( \langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle \);
4. \( \lambda \langle u, v \rangle = \langle \lambda u, v \rangle \);
5. \( \langle u + v, u + v \rangle + \langle u - v, u - v \rangle = 2 \langle u, u \rangle + 2 \langle v, v \rangle \)

For every \( v \in V \), we define the norm of \( v \), denoted by \( ||v|| \), to be \( ||v|| = \sqrt{\langle v, v \rangle} \).

**Proposition 1.13.** Let \( u, v \in V \) and \( \lambda \in \mathbb{R} \). Then

1. \( ||\lambda v|| = |\lambda|||v|| \);
2. \( ||\langle u, v \rangle|| \leq ||u|| \cdot ||v|| \);
3. \( ||u + v|| \leq ||u|| + ||v|| \) and the equality holds when \( u = \lambda v \) for some \( \lambda \in \mathbb{R} \).
Proof. (1) is trivial. For (2), which is usually called Schwartz inequality, since
\begin{equation}
\langle u + \lambda v, u + \lambda v \rangle = (\|v\|^2)\lambda^2 + (2\langle u, v \rangle)\lambda + \|u\|^2 \geq 0
\end{equation}
for any \(u, v \in V\) and \(\lambda \in \mathbb{R}\),
\begin{equation}
(2\langle u, v \rangle)^2 - 4\|u\|^2 \cdot \|v\|^2 \leq 0
\end{equation}
and (2) follows. (3) follows from (2) directly. \(\square\)

For any two points \(p, q \in \mathbb{R}^n\), we define the distance between \(p\) and \(q\) as
\begin{equation}
d(p, q) = \|\overrightarrow{op} - \overrightarrow{oq}\|
\end{equation}
where \(o\) is the origin. It is easy to check

Proposition 1.14. Let \(x, y, z \in \mathbb{R}^n\).

(1) (Triangle Inequality) \(d(x, y) + d(y, z) \geq d(z, x)\);
(2) \(d(x, y) \geq 0\) and the equality holds if and only if \(x = y\).

\(d(\cdot, \cdot)\) is called a metric on and this makes \(\mathbb{R}^n\) a metric space. It induces a topology on \(\mathbb{R}^n\).

For \(x \in \mathbb{R}^n\) and \(r > 0\), we define
\begin{equation}
B(x, r) = \{y \in \mathbb{R}^n : d(y, x) < r\}
\end{equation}
to be the open ball with center \(x\) and radius \(r\).

Let \(S \subset \mathbb{R}^n\). A point \(x \in S\) is an interior point of \(S\) if \(B(x, r) \subset S\) for some \(r > 0\). The interior of \(S\), denoted by \(\text{Int}(S)\), is the subset of \(S\) consisting of all interior points of \(S\). \(S\) is open if \(S = \text{Int}(S)\), i.e., every point of \(S\) is an interior point of \(S\).

Proposition 1.15. \(\text{Int}(S) \subset S\) is open.

(1) The union of any collection of open sets is open.
(2) The intersection of any finitely many open sets is open.
(3) \(\text{Int}(S)\) is the union of all open subsets that are contained in \(S\).

Proof. For (1), let \(\{S_\alpha\}_{\alpha \in I}\) be a collection of open sets. For any point \(x \in \bigcup S_\alpha\), \(x \in S_\alpha\) for some \(\alpha \in I\). Since \(S_\alpha\) is open, there exists \(r > 0\) such that \(B(x, r) \subset S_\alpha\) and \(B(x, r) \subset \bigcup S_\alpha\). Consequently, \(x\) is an interior point of \(\bigcup S_\alpha\) and \(\bigcup S_\alpha\) is open.

For (2), it suffices to prove the intersection of two open sets is open. Let \(S_1\) and \(S_2\) be two open sets. Let \(x \in S_1 \cap S_2\). Since \(S_1\) and \(S_2\) are open, there exist \(r_1, r_2 > 0\) such that \(B(x, r_1) \subset S_1\) and \(B(x, r_2) \subset S_2\). Let \(r = \min(r_1, r_2)\). Then \(B(x, r) \subset S_1 \cap S_2\) and \(x\) is a point in \(S_1 \cap S_2\), and hence \(S_1 \cap S_2\) is open.

For (3), it is obvious that \(\text{Int}(S) \subset S\) is open. It suffices to show that any open set \(T \subset S\) is contained in \(\text{Int}(S)\). Let \(x \in T\). Since \(T\) is open, there exists \(r > 0\) such that \(B(x, r) \subset T \subset S\). So \(x\) is an interior point of \(S\) and \(x \in \text{Int}(S)\). Therefore, \(T \subset \text{Int}(S)\). \(\square\)
A set $S$ is called closed if the complement of $S$ is open. The closure of $S$, denoted by $\text{cl}(S)$, is the intersection of all closed sets that contain $S$.

**Proposition 1.16.**

1. The intersection of any collection of closed sets is closed.
2. The union of any finitely many closed sets is closed.
3. $x \in \text{cl}(S)$ if and only if $B(x, r) \cap S \neq \emptyset$ for all $r > 0$.
4. $\text{cl}(S^c) = \text{Int}(S)^c$.

**Proof.** (1) and (2) follow directly from Proposition 1.15.

For (3), let $x \in \text{cl}(S)$. If $B(x, r) \cap S = \emptyset$ for some $r > 0$, $B(x, r)^c \supset S$ and then $x \in B(x, r)^c$ by the definition of cl$(S)$: contradiction. Therefore, $B(x, r) \cap S \neq \emptyset$ for all $r > 0$. On the other hand, let $x$ be a point with the property that $B(x, r) \cap S \neq \emptyset$ for all $r > 0$. To show that $x \in \text{cl}(S)$, it suffices to show that $x \in T$ for every closed set $T \supset S$. Suppose that $x \notin T$. Then $x \in T^c$. Since $T^c$ is open, $B(x, r) \subset T^c$ for some $r > 0$. Hence $B(x, r) \cap T = \emptyset$ and $B(x, r) \cap S = \emptyset$. Contradiction.

For (4), first notice that $S^c \subset \text{Int}(S)^c$. Since $\text{Int}(S)^c$ is closed, $\text{cl}(S^c) \subset \text{Int}(S)^c$. Let $x \in \text{Int}(S)^c$. If $x \notin \text{cl}(S^c)$, then $B(x, r) \cap S^c = \emptyset$ for some $r > 0$. Consequently, $B(x, r) \subset S$ and $x \in \text{Int}(S)$. Contradiction. Therefore, $x \in \text{cl}(S^c)$ and (4) follows. □

The boundary of $S$, denoted by $\text{bd}(S)$, is the difference $\text{cl}(S) - \text{Int}(S)$. It is not hard to see that $x \in \text{bd}(S)$ if $B(x, r) \cap S \neq \emptyset$ and $B(x, r) \cap S^c \neq \emptyset$ for any $r > 0$.

Let $S$ be a subset of $\mathbb{R}^n$. $S$ is contained in its affine hull $\text{aff}(S) \cong \mathbb{R}^k$. The relative interior of $S$, denoted by $\text{relint}(S)$, is the interior of $S$ as a subset of $\mathbb{R}^k$.

A map $\mathbb{R}^n \to \mathbb{R}^m$ is called continuous if $f^{-1}(U)$ is open for any open set $U \subset \mathbb{R}^m$.

A set $S \subset \mathbb{R}^n$ is bounded if $S \subset B(o, R)$ for some $R > 0$. A set $S \subset \mathbb{R}^n$ is compact if it is closed and bounded.

**Theorem 1.17.** The images of compact sets under a continuous map are compact.

We will not prove this theorem here. Interested students may check any standard real analysis book for a proof.

**Theorem 1.18 (Extreme Value Theorem).** Let $S$ be a compact set and $f : S \to \mathbb{R}$ be a continuous function on $S$. Then $f(x)$ achieves maximum and minimum on $S$.

Again, one may find a proof for this theorem in any standard real analysis book.

A set $S$ is connected if there do not exist two open sets $A$ and $B$ such that $A \cap S \neq \emptyset$, $B \cap S \neq \emptyset$, $A \cap B = \emptyset$ and $S \subset A \cup B$. 

1. EUCLIDEAN GEOMETRY 11
2. Convex Sets

2.1. Definitions. Let \( x, y \in \mathbb{R}^n \). We use the notation \( \overline{xy} \) to denote the line segment
\[
\overline{xy} = \{ \alpha x + \beta y : \alpha, \beta \geq 0, \alpha + \beta = 1 \}
\]
A set \( S \subset \mathbb{R}^n \) is convex if \( \overline{xy} \subset S \) for any \( x, y \in S \).

We call a vector \( x \) a convex combination of \( x_1, x_2, ..., x_m \) if there exists \( \lambda_1, \lambda_2, ..., \lambda_m \geq 0 \) such that \( \lambda_1 + \lambda_2 + ... + \lambda_m = 1 \) and \( x = \lambda_1 x_1 + \lambda_2 x_2 + ... + \lambda_m x_m \). Obviously, \( \overline{xy} \) is the set consisting of all convex combinations of the two points \( x \) and \( y \). So \( S \) is convex if it contains all convex combinations of any two points in \( S \). Actually, this is true for any finitely many points in \( S \).

That is, we have

**Proposition 2.1.** \( S \) is convex if and only if it contains all convex combinations of any finitely many points in \( S \).

The proof of the above proposition goes exactly like that of Proposition 1.9. The following is obvious.

We will leave the proofs of the following two statements as exercises.

**Proposition 2.2.** The intersection of any collection of convex sets is convex.

**Proposition 2.3.** Let \( f : V \to W \) be a linear transformation between two real vector spaces \( V \) and \( W \). Then \( f(S) \) is convex for any convex set \( S \subset V \) and \( f^{-1}(T) \) is convex for any convex set \( T \subset W \).

The convex hull of a set \( S \), denoted by \( \text{conv}(S) \), is the intersection of all convex sets which contain \( S \).

**Proposition 2.4.** The convex hull of \( S \) consists precisely of all convex combinations of points of \( S \). That is,
\[
\text{conv}(S) = T = \{ \lambda_1 x_1 + \lambda_2 x_2 + ... + \lambda_m x_m : \lambda_i \geq 0, \sum_{i=1}^{m} \lambda_i = 1, x_i \in S \text{ for } i = 1, 2, ..., m \}
\]

The proof of the above proposition goes exactly like that of Proposition 1.11, which we will not repeat here.

Next we will show the interior and closure of a convex set are still convex.

**Proposition 2.5.** Let \( S \) be a convex set. If \( x \in \text{Int}(S) \) and \( y \in S \), then \( \text{relint} \overline{xy} \subset \text{Int}(S) \). Consequently, \( \text{Int}(S) \) is convex.

**Proof.** Note that
\[
\text{relint} \overline{xy} = \{ \alpha x + \beta y : \alpha, \beta > 0, \alpha + \beta = 1 \}
\]
Let \( z = \alpha x + \beta y \in \text{relint} \overline{xy} \). Since \( x \in \text{Int}(S) \), \( B(x, r) \subset S \) for some \( r > 0 \). Since \( S \) is convex, \( aB(x, r) + \beta y \subset S \). Note that
\[
\alpha B(x, r) + \beta y = B(\alpha x, \alpha r) + \beta y = B(\alpha x + \beta y, \alpha r) = B(z, \alpha r)
\]
Therefore, \( z \in \operatorname{Int}(S) \) and \( \overline{\operatorname{relint} H} \subset \operatorname{Int}(S) \). 

**Proposition 2.6.** Let \( S \) be a convex set. Then \( \overline{\operatorname{Cl}}(S) \) is convex.

**Proof.** Let \( x, y \in \overline{\operatorname{Cl}}(S) \). To show \( \overline{\operatorname{Cl}}(S) \) is convex, it suffices to show \( \alpha x + \beta y \in \overline{\operatorname{Cl}}(S) \) for all \( \alpha, \beta \geq 0 \) with \( \alpha + \beta = 1 \). Since \( x \in \overline{\operatorname{Cl}}(S) \), there exists a sequence \( \{x_n \in S\} \) such that \( d(x, x_n) \to 0 \) as \( n \to \infty \). Similarly, there exists a sequence \( \{y_n \in S\} \) such that \( d(y, y_n) \to 0 \) as \( n \to \infty \). To show that \( z = \alpha x + \beta y \in \overline{\operatorname{Cl}}(S) \), it suffices to show that there exists a sequence \( \{z_n \in S\} \) such that \( d(z, z_n) \to 0 \) as \( n \to \infty \). Let \( z_n = \alpha x_n + \beta y_n \). Obviously, \( z_n \in S \) since \( x_n, y_n \in S \) and \( S \) is convex. Then

\[
d(z, z_n) = ||z - z_n|| = ||\alpha(x - x_n) + \beta(y - y_n)||
\]

\[
\leq \alpha||x - x_n|| + \beta||y - y_n|| = \alpha d(x, x_n) + \beta d(y, y_n)
\]

Therefore, \( \lim_{n \to \infty} d(z, z_n) = 0 \) because \( \lim_{n \to \infty} d(x, x_n) = \lim_{n \to \infty} d(y, y_n) = 0 \). 

**Proposition 2.7.** If \( S \) is an open set, then \( \overline{\operatorname{Cl}}(S) \) is also open.

**Proof.** It suffices to show that \( \overline{\operatorname{Cl}}(S) \subset \operatorname{Int(\operatorname{Cl}(S))} \). Since \( S \subset \overline{\operatorname{Cl}}(S) \), \( \operatorname{Int}(S) \subset \operatorname{Int(\operatorname{Cl}(S))} \). Since \( S \) is open, \( S \subset \operatorname{Int(\operatorname{Cl}(S))} \). By Proposition 2.5, \( \operatorname{Int}(\overline{\operatorname{Cl}}(S)) \) is convex. Therefore, \( \overline{\operatorname{Cl}}(S) \subset \operatorname{Int(\operatorname{Cl}(S))} \).

It is natural to expect the convex hull of a closed set is also closed. However, this is not true. For example, let \( f : \mathbb{R} \to \mathbb{R} \) be an arbitrary positive continuous function satisfying

\[
\lim_{x \to -\infty} f(x) = \lim_{x \to -\infty} f(x) = 0
\]

and \( S = \{(x, y) : y \geq f(x)\} \). Since \( f(x) \) is continuous, \( S \) is closed but we claim that \( \overline{\operatorname{Cl}}(S) = \{(x, y) : y > 0\} \), which is not closed. Let \( r = (x_0, y_0) \in \mathbb{R}^2 \) with \( y_0 > 0 \). Since \( f(x) \to 0 \) as \( x \to -\infty \), there exists \( x_1 < x_0 \) such that \( f(x_1) < y_0 \). Similarly, there exists \( x_2 > x_0 \) such that \( f(x_2) < y_0 \). Let \( p = (x_1, y_0) \) and \( q = (x_2, y_0) \). Since \( y_0 \geq f(x_i) \) for \( i = 1, 2 \), \( p, q \in S \). So \( \overline{\operatorname{pq}} \subset \overline{\operatorname{Cl}}(S) \). Obviously, \( r \in \overline{\operatorname{pq}} \) and therefore \( r \in \overline{\operatorname{Cl}}(S) \). Hence \( H = \{(x, y) : y > 0\} \subset \overline{\operatorname{Cl}}(S) \). On the other hand, \( S \subset H \) and \( H \) is convex. Therefore, \( H = \overline{\operatorname{Cl}}(S) \).

Therefore, the convex hull of a closed set \( S \subset \mathbb{R}^n \) is not necessarily convex. However, if we further assume that \( S \) is bounded, i.e., \( S \) is compact, \( \overline{\operatorname{Cl}}(S) \) is compact. The proof of this theorem requires Carathéodory’s theorem.

### 2.2. Carathéodory’s Theorem.

**Theorem 2.8 (Carathéodory).** If \( S \) is nonempty subset of \( \mathbb{R}^n \), then every \( x \) in \( \overline{\operatorname{Cl}}(S) \) can be expressed as a convex combination of \( n + 1 \) or fewer points of \( S \).
1. BASICS

**Proof.** Let $m$ be the smallest number such that $x$ is the convex combination of $m$ points. Then there exist $x_1, x_2, \ldots, x_m \in S$ and $\lambda_1, \lambda_2, \ldots, \lambda_m \geq 0$ such that

\[(2.7) \quad \sum_{i=1}^{m} \lambda_i = 1 \quad \text{and} \quad x = \sum_{i=1}^{m} \lambda_i x_i\]

Due to our choice of $m$, $\lambda_i > 0$ for all $i = 1, 2, \ldots, m$.

Suppose that $m > n + 1$. By Corollary 1.8, $x_1, x_2, \ldots, x_m$ are affinely dependent, i.e., there exist $\gamma_1, \gamma_2, \ldots, \gamma_m$, not all zero, such that

\[(2.8) \quad \sum_{i=1}^{m} \gamma_i = 0 \quad \text{and} \quad \sum_{i=1}^{m} \gamma_i x_i = 0\]

Let $I = \{i : \gamma_i > 0\} \subset \{1, 2, \ldots, m\}$. Since $\gamma_1, \gamma_2, \ldots, \gamma_m$ are not all zero, $I \neq \emptyset$. Let

\[(2.9) \quad a = \min_{i \in I} \lambda_i \gamma_i\]

and let $\alpha_i = \lambda_i - a \gamma_i$. If $i \in I$, then $\alpha_i \geq 0$ since $a \geq \lambda_i / \gamma_i$; if $i \notin I$, $\gamma_i \leq 0$ and hence $\alpha_i \geq 0$. Therefore, $\alpha_i \geq 0$ for all $i = 1, 2, \ldots, m$.

By (2.8),

\[(2.10) \quad \sum_{i=1}^{m} \alpha_i = \sum_{i=1}^{m} \lambda_i = 1 \quad \text{and} \quad \sum_{i=1}^{m} \alpha_i x_i = \sum_{i=1}^{m} \lambda_i x_i = x\]

Due to our choice of $a$, there is at least one $i \in I$ such that $a = \lambda_i / \gamma_i$. Correspondingly, $\alpha_i = 0$. Therefore by (2.10), $x$ is a convex combination of $x_1, x_2, \ldots, \hat{x_i}, \ldots, x_m$. Contradiction.

**Proposition 2.9.** If $S \subset \mathbb{R}^n$ is a compact set, then $\text{conv}(S)$ is also compact.

**Proof.** By Caratheodory’s theorem, every point $x \in \text{conv}(S)$ is the convex combination of some $n + 1$ points $x_1, x_2, \ldots, x_{n+1}$ in $S$. Let $f : (\mathbb{R}^n)^{n+1} \times \mathbb{R}^{n+1} : \mathbb{R}^n$ be the map:

\[(2.11) \quad f(x_1, x_2, \ldots, x_{n+1}, \lambda_1, \lambda_2, \ldots, \lambda_{n+1}) = \lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_{n+1} x_{n+1}\]

where $x_i \in \mathbb{R}^n$ and $\lambda_i \in \mathbb{R}$. Let

\[(2.12) \quad D = \{(\lambda_1, \lambda_2, \ldots, \lambda_{n+1}) : \lambda_1, \lambda_2, \ldots, \lambda_{n+1} \geq 0, \lambda_1 + \lambda_2 + \ldots + \lambda_{n+1} = 1\}\]

Then $f$ is continuous and $D$ is compact.

By Caratheodory’s theorem, $f(S^{n+1} \times D) = \text{conv}(S)$. Since $S^{n+1} \times D$ is compact, $\text{conv}(S)$ is compact.

**Proposition 2.10.** Let $S \subset \mathbb{R}^n$ be a convex set. If $S \neq \emptyset$, then $\text{relint}(S) \neq \emptyset$. 


2. CONVEX SETS

Proof. Consider $S$ as a convex subset of $\text{aff}(S) = \mathbb{R}^k$. We will show that $\text{Int}(S) \neq \emptyset$.

Since $\text{aff}(S) = \mathbb{R}^k$, there exist $x_1, x_2, ..., x_{k+1} \in S$ which are affinely independent. Let $D \subset \mathbb{R}^{k+1}$ be the subset

$$D = \{ (\lambda_1, \lambda_2, ..., \lambda_{k+1}) : \lambda_1 + \lambda_2 + ... + \lambda_{k+1} = 1 \}$$

Let $f : D \rightarrow \mathbb{R}^k$ be the map

$$f(\lambda_1, \lambda_2, ..., \lambda_{k+1}) = \lambda_1 x_1 + \lambda_2 x_2 + ... + \lambda_{k+1} x_{k+1}$$

$f$ is obviously continuous and onto. Also we claim that $f$ is one-to-one. Otherwise, there exist $(\lambda'_1, \lambda'_2, ..., \lambda'_{k+1}) \neq (\lambda_1, \lambda_2, ..., \lambda_{k+1}) \in D$ such that $f(\lambda_1, \lambda_2, ..., \lambda_{k+1}) = f(\lambda'_1, \lambda'_2, ..., \lambda'_{k+1})$, which implies

$$(\lambda_1 - \lambda'_1)x_1 + (\lambda_2 - \lambda'_2)x_2 + ... + (\lambda_{k+1} - \lambda'_{k+1}) x_{k+1} = 0$$

and hence $x_1, x_2, ..., x_{k+1}$ are affinely dependent. Contradiction. Hence $f$ is one-to-one. Let $g = f^{-1} : \mathbb{R}^k \rightarrow D$ be the inverse of $f$ and let

$$g(z_1, z_2, ..., z_k) = (g_1(z_1, z_2, ..., z_k), g_2(z_1, z_2, ..., z_k), ..., g_{k+1}(z_1, z_2, ..., z_k))$$

where $g_i$ are functions from $\mathbb{R}^k \rightarrow \mathbb{R}$.

More explicitly, let $x_j = (a_{1j}, a_{2j}, ..., a_{kj})$ and $A$ be the $(k+1) \times (k+1)$ matrix

$$A = \begin{bmatrix}
    a_{11} & a_{12} & \cdots & a_{1,k+1} \\
    a_{21} & a_{22} & \cdots & a_{2,k+1} \\
    \vdots & \vdots & \ddots & \vdots \\
    a_{k1} & a_{k2} & \cdots & a_{k,k+1} \\
    1 & 1 & \cdots & 1
\end{bmatrix}$$

Then

$$\begin{bmatrix}
    g_1(z_1, z_2, ..., z_k) \\
    g_2(z_1, z_2, ..., z_k) \\
    \vdots \\
    g_{k+1}(z_1, z_2, ..., z_k)
\end{bmatrix} = A^{-1} \begin{bmatrix}
    z_1 \\
    z_2 \\
    \vdots \\
    z_k \\
    1
\end{bmatrix}$$

Therefore, each $g_i(z_1, z_2, ..., z_k)$ is a (nonhomogeneous) linear function in $z_1, z_2, ..., z_k$. Therefore, $g_i$ is continuous and $g$ is continuous.

Let

$$x_0 = \frac{1}{k+1} (x_1 + x_2 + ... + x_{k+1}) \in S$$

Then

$$y_0 = g(x_0) = \left( \frac{1}{k+1}, \frac{1}{k+1}, \ldots, \frac{1}{k+1} \right)$$
Choose an arbitrary \( r \) with \( 0 < r < 1/(k + 1) \). Since \( g \) is continuous, \( U = g^{-1}(B(y_0, r) \cap D) \) is an open set in \( \mathbb{R}^k \) that contains the point \( x_0 \). Note

\[
U = \{ \lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_{k+1} x_{k+1} : (\lambda_1, \lambda_2, \ldots, \lambda_{k+1}) \in B(y_0, r) \cap D \}
\]

Since \( 0 < r < 1/(k + 1) \), \( \lambda_1, \lambda_2, \ldots, \lambda_{k+1} > 0 \) for every \( (\lambda_1, \lambda_2, \ldots, \lambda_{k+1}) \in B(y_0, r) \). Therefore, \( U \subset S \) and \( x_0 \in \text{Int}(S) \). \( \square \)
CHAPTER 2

Some Selected Topics in Convex Geometry

1. Helly’s Theorem

Theorem 1.1. Let \( S_1, S_2, \ldots, S_m \) be \( m \geq n + 1 \) convex sets in \( \mathbb{R}^n \). If every \( n+1 \) sets among \( S_1, S_2, \ldots, S_m \) have nonempty intersection, \( S_1 \cap S_2 \cap \ldots \cap S_m \neq \emptyset \).

Proof. We prove by induction on \( m \). It is trivial for \( m = n + 1 \). Suppose that it holds for \( m < l \), where \( l > n + 1 \). We want to show it holds for \( m = l \).

By induction hypothesis, any \( m-1 \) sets among \( S_1, S_2, \ldots, S_m \) have nonempty intersection. That is, if we let

\[
T_k = S_1 \cap S_2 \cap \ldots \cap \hat{S}_k \cap \ldots \cap S_m
\]

for every \( k = 1, 2, \ldots, m \). Since \( T_k \neq \emptyset \), we choose (arbitrarily) a point \( x_k \in T_k \). Since \( m \geq n + 2 \), \( x_1, x_2, \ldots, x_m \) are affinely dependent. That is, there exist \( \lambda_1, \lambda_2, \ldots, \lambda_m \in \mathbb{R} \), not all zero, such that

\[
\lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_m x_m = 0
\]

and

\[
\lambda_1 + \lambda_2 + \ldots + \lambda_m = 0
\]

Obviously, we may rewrite (1.2) as

\[
\sum_{\lambda_i > 0} \lambda_i x_i = -\sum_{\lambda_j \leq 0} \lambda_j x_j
\]

By (1.3),

\[
\lambda = \sum_{\lambda_i > 0} \lambda_i = -\sum_{\lambda_j \leq 0} \lambda_j
\]

Since \( \lambda_1, \lambda_2, \ldots, \lambda_m \) are not all zero, \( \lambda > 0 \). Hence

\[
x = \sum_{\lambda_i > 0} \left( \frac{\lambda_i}{\lambda} \right) x_i = \sum_{\lambda_j \leq 0} -\left( \frac{\lambda_j}{\lambda} \right) x_j
\]

Therefore, \( x \) is a convex combination of \( \{ x_i : \lambda_i > 0 \} \) and it is also a convex combination of \( \{ x_j : \lambda_j \leq 0 \} \). Let \( I = \{ i : \lambda_i > 0 \} \) and \( J = \{ j : \lambda_j \leq 0 \} \). Obviously, \( I \cup J = \{ 1, 2, \ldots, m \} \) and \( I \cap J = \emptyset \). Since \( x_i \in S_j \) for any \( i \neq j \),
$x_i \in \cap_{j \in J} S_j$ for every $i \in I$. Since $\cap_{j \in J} S_j$ is convex, $x \in \cap_{i \in I} S_i$. By the same argument, we see that $x \in \cap_{i \in I} S_i$. Therefore,

(1.7) \hspace{1cm} x \in (\bigcap_{i \in I} S_i) \cap (\bigcap_{j \in J} S_j) = \bigcap_{k=1}^m S_k

Therefore, $\cap_{k=1}^m S_k \neq \emptyset$. \hfill \Box

Obviously, the number $n + 1$ in the theorem is optimal. It is easy to find, for example, three convex sets in $\mathbb{R}^2$ such that any two of them have nonempty intersection but all three of them have empty intersection.

2. Introduction to Linear Programming

Let us consider the following question:

**Question 2.1.** Given points $x, x_1, x_2, \ldots, x_m \in \mathbb{R}^n$, determine whether $x \in \text{conv}\{x_1, x_2, \ldots, x_m\}$.

Here we assume that $\dim \text{aff}\{x_1, x_2, \ldots, x_m\} = n$; otherwise, we can restrict ourselves to $\mathbb{R}^k = \text{aff}\{x_1, x_2, \ldots, x_m\}$.

The case $m = n + 1$ is easy to treat. Since $x_1, x_2, \ldots, x_{n+1}$ are affinely independent, $x$ can be written as an affine combination of $x_1, x_2, \ldots, x_{n+1}$ in a unique way:

(2.1) \hspace{1cm} x = \lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_{n+1} x_m

where

(2.2) \hspace{1cm} \lambda_1 + \lambda_2 + \ldots + \lambda_m = 1.

Therefore, $x \in \text{conv}\{x_1, x_2, \ldots, x_{n+1}\}$ if and only if $\lambda_i \geq 0$ for all $i = 1, 2, \ldots, n+1$.

However, the situation is more complicated when $m > n + 1$. We let $x_i = (a_{i1}, a_{i2}, \ldots, a_{in})$ and $x = (b_1, b_2, \ldots, b_n)$. Then (2.1) and (2.2) expand to

(2.3) \hspace{1cm} \begin{align*}
  a_{11} \lambda_1 &+ a_{12} \lambda_2 + \ldots + a_{1m} \lambda_m = b_1 \\
  a_{21} \lambda_1 &+ a_{22} \lambda_2 + \ldots + a_{2m} \lambda_m = b_2 \\
  \vdots & \vdots \ldots \vdots \vdots \\
  a_{n1} \lambda_1 &+ a_{n2} \lambda_2 + \ldots + a_{nm} \lambda_m = b_n \\
  \lambda_1 &+ \lambda_2 + \ldots + \lambda_m = 1
\end{align*}

with $\lambda_1, \lambda_2, \ldots, \lambda_m \geq 0$. Obviously, $x \in \text{conv}\{x_1, x_2, \ldots, x_{n+1}\}$ if and only if (2.3) has a solution. Here we may assume $b_i \geq 0$ for all $i$; otherwise, if $b_i < 0$, we replace the $i$-th equation by

(2.4) \hspace{1cm} -a_{i1} \lambda_1 - a_{i2} \lambda_2 - \ldots - a_{im} \lambda_m = -b_i

For reasons that will be clear later, we add $n$ “auxiliary” variables $\lambda_{m+1}, \lambda_{m+2}, \ldots, \lambda_{m+n}, \lambda_{m+n+1}$. We consider the following problem:
Question 2.2. Minimize the function

\[ f(\lambda_1, \lambda_2, ..., \lambda_m, \lambda_{m+1}, \lambda_{m+2}, ..., \lambda_{m+n+1}) \]

\[ = \lambda_{m+1} + \lambda_{m+2} + ... + \lambda_{m+n+1} \]

with \( \lambda_i \geq 0 \) (i = 1, 2, ..., m, ..., m + n + 1) satisfying

\[
\begin{align*}
\sum_{j=1}^{m} a_{1j} \lambda_j + \lambda_{m+1} &= b_1 \\
\sum_{j=1}^{m} a_{2j} \lambda_j + \lambda_{m+2} &= b_2 \\
\sum_{j=1}^{m} a_{nj} \lambda_j + \lambda_{m+n} &= b_n \\
\sum_{j=1}^{m} \lambda_j + \lambda_{m+n+1} &= 1
\end{align*}
\]

Lemma 2.3. \( x \in \text{conv}\{x_1, x_2, ..., x_m\} \) if and only if \( f_{\text{min}} = 0 \) in Question 2.2.

Question 2.2 is a typical problem in so-called Linear Programming (LP). One of the many algorithms to solve a LP problem is Simplex Method. We will use the following example to demonstrate the algorithm.

Example 2.1. Let \( x_1 = (1, 0), \ x_2 = (0, 1), \ x_3 = (1, 1) \) and \( x_4 = (-2, -2) \). Determine whether \( x = (-1, -1) \in \text{conv}\{x_1, x_2, x_3, x_4\} \).

First we need to formulate the problem as in Question 2.2. Obviously, \( x \in \text{conv}\{x_1, x_2, x_3, x_4\} \) if and only if the system of linear equations

\[
\begin{align*}
\lambda_1 + \lambda_3 - 2\lambda_4 &= -1 \\
\lambda_2 + \lambda_3 - 2\lambda_4 &= -1 \\
\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 &= 1
\end{align*}
\]

has a solution with \( \lambda_i \geq 0 \) for \( i = 1, 2, 3, 4 \).

It is a requirement of simplex method that \( b_i \geq 0 \). Therefore, we change (2.7) to the form

\[
\begin{align*}
-\lambda_1 - \lambda_3 + 2\lambda_4 &= 1 \\
-\lambda_2 - \lambda_3 + 2\lambda_4 &= 1 \\
\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 &= 1
\end{align*}
\]

By adding three auxiliary variables \( \lambda_5, \lambda_6, \lambda_7 \), we put it into a LP problem:

(*) Minimize \( f(\lambda_1, ..., \lambda_7) = \lambda_5 + \lambda_6 + \lambda_7 \) with \( \lambda_i \geq 0 \) satisfying

\[
\begin{align*}
-\lambda_1 - \lambda_3 + 2\lambda_4 &= 1 \\
-\lambda_2 - \lambda_3 + 2\lambda_4 &= 1 \\
\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 &= 1 \\
\lambda_5 + \lambda_6 + \lambda_7 &= 1
\end{align*}
\]

In order to apply simplex method to (*), we need first to put the objective function \( f \) in terms of \( \lambda_1, \lambda_2, \lambda_3, \lambda_4 \) (using (2.9))

\[
\begin{align*}
f(\lambda_1, ..., \lambda_7) &= \lambda_3 - 5\lambda_4 + 3
\end{align*}
\]

We put the coefficients of (2.9) and (2.10) into a table called simplex tableau:
2. SOME SELECTED TOPICS IN CONVEX GEOMETRY

\[
\begin{pmatrix}
-1 & 0 & -1 & 2 & 1 & 0 & 0 & \vdots & 1 \\
0 & -1 & -1 & 2 & 0 & 1 & 0 & \vdots & 1 \\
1 & 1 & 1 & 1 & 0 & 0 & 1 & \vdots & 1 \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
0 & 0 & 1 & -5 & 0 & 0 & 0 & \vdots & -3 \\
\end{pmatrix}
\]

(2.11)

Note that the last row \((c_1, c_2, ..., c_{m+n+1}, -A)\) corresponds to the objective function

\[
f(\lambda_1, \lambda_2, ..., \lambda_{m+n+1}) = c_1\lambda_1 + c_2\lambda_2 + \cdots + c_{m+n+1}\lambda_{m+n+1} + A
\]

(2.12)

The algorithm goes as follows:

1. If the entries of the last row except the rightmost entry, i.e., \(c_j \geq 0\) for all \(j\), stop and we are done. The minimum of \(f\) is given by \(-A\), where \(-A\) is the rightmost entry of the last row. And \(x \in \text{conv}\{x_1, x_2, ..., x_m\}\) if and only if \(A = 0\), i.e., the rightmost entry of the last row vanishes.

2. If one of \(c_j < 0\), we look at the \(j\)-th column. We choose the entry \(a_{kj}\) (called pivoting term) such that \(a_{kj} > 0\) and

\[
\frac{b_k}{a_{kj}} = \min \left\{ \frac{b_i}{a_{ij}} : a_{ij} > 0 \right\}
\]

(2.13)

Then we use \(a_{kj}\) to eliminate all the other entries of the \(j\)-the column by row operations.

3. Repeat step (2) until (1) happens.

Now let's carry it out for our example: (the pivoting term of each step is boxed)
Therefore, $f_{\min} = 0$ and $x \in \text{conv}\{x_1, x_2, x_3, x_4\}$. In addition, the final tableau also tells us how to write $x$ as a convex combination of $x_1, x_2, x_3, x_4$.

Note that if we let $\lambda_3 = \lambda_5 = \lambda_6 = \lambda_7 = 0$, then $2\lambda_4 = 6/5$, $\lambda_1 = 1/5$ and $(5/2)\lambda_2 = 1/2$. Consequently, $x = (1/5)x_1 + (1/5)x_2 + (3/5)x_4$.

3. Convex Functions

Given a function $f(x)$, we consider the region $S = \{(x, y) : y \geq f(x)\} \subset \mathbb{R}^2$. There is a very useful criterion on the convexity of $S$.

**Theorem 3.1.** Suppose that $f(x)$ is twice differentiable and $f''(x) \geq 0$ for all $a \leq x \leq b$. Then the region $S = \{(x, y) : y \geq f(x), a \leq x \leq b\}$ is convex.

An easy corollary of the above theorem is the following:
Corollary 3.2. Suppose that \( f(x) \) is twice differentiable and \( f''(x) \leq 0 \) for all \( a \leq x \leq b \). Then the region \( S = \{(x, y) : y \leq f(x), a \leq x \leq b\} \) is convex.

Proof. Since \( f''(x) \leq 0 \), \( -f''(x) \geq 0 \) for all \( a \leq x \leq b \). Then \( S' = \{(x, y) : y \geq -f(x)\} \) is convex by the above theorem. Geometrically, \( S \) is the reflection of \( S' \) with respect to the \( x \)-axis. More precisely, let \( g : \mathbb{R}^2 \to \mathbb{R}^2 \) be the linear map given by \( g(x, y) = (x, -y) \). Then \( S = g(S') \). Since the images of convex sets under linear maps are convex, \( S \) is convex. \( \square \)

The proof of Theorem 3.1 is elementary. It uses nothing more than Mean Value Theorem (MVT).

Let \( p = (x_p, y_p) \) and \( q = (x_q, y_q) \) be two points in \( S \). WLOG, assume that \( x_p < x_q \). Since \( p, q \in S \), \( y_p \geq f(x_p) \) and \( y_q \geq f(x_q) \). Suppose that \( \overline{pq} \not\in S \). Then there exists a point \( r = (x_r, y_r) \in \overline{pq} \) such that \( r \not\in S \), i.e., \( y_r < f(x_r) \). By MVT, there exists \( x_u \in (x_p, x_r) \) such that

\[
(3.1) \quad f'(x_u) = \frac{f(x_r) - f(x_p)}{x_r - x_p}
\]

and there exists \( x_v \in (x_r, x_q) \) such that

\[
(3.2) \quad f'(x_v) = \frac{f(x_q) - f(x_r)}{x_q - x_r}
\]

Since \( y_r < f(x_r) \) and \( y_v \geq f(x_v) \),

\[
(3.3) \quad \frac{f(x_r) - f(x_p)}{x_r - x_p} \geq \frac{y_r - y_p}{x_r - x_p}
\]

By the same reason,

\[
(3.4) \quad \frac{f(x_q) - f(x_r)}{x_q - x_r} \leq \frac{y_q - y_r}{x_q - x_r}
\]

Since \( p, q, r \) are collinear,

\[
(3.5) \quad \frac{y_r - y_p}{x_r - x_p} = \frac{y_q - y_r}{x_q - x_r}
\]

Therefore,

\[
(3.6) \quad \frac{f(x_r) - f(x_p)}{x_r - x_p} \geq \frac{f(x_q) - f(x_r)}{x_q - x_r}
\]

That is, \( f'(x_u) > f'(x_v) \). This contradicts the fact that \( f''(x) \geq 0 \) and \( f'(x) \) is nondecreasing on \([a, b]\).

Theorem 3.1 can be generalized to dimension greater than 2. We will state the following without proof.

Theorem 3.3. Let \( f(x_1, x_2, \ldots, x_n) \) be a twice differentiable function in \( n \) variables. Suppose that the \( n \times n \) matrix (called the Hessian of \( f \))

\[
(3.7) \quad H = \left[ \frac{\partial^2 f}{\partial x_i \partial x_j} \right]_{n \times n}
\]
is positive definite for every \((x_1, x_2, ..., x_n)\) in a convex set \(D \subset \mathbb{R}^n\). Then the region
\[
(3.8) \quad S = \{ (x_1, x_2, ..., x_{n+1}) : x_{n+1} \geq f(x_1, x_2, ..., x_n), (x_1, x_2, ..., x_n) \in D \}
\]
in \(\mathbb{R}^{n+1}\) is convex.
Bibliography

CHAPTER 1. Basics. 1. Euclidean Geometry. 1.1. Vector spaces. Let $\mathbb{R}$ be the set of real numbers. A set $S \subset \mathbb{R}^n$ is convex if $xy \subset S$ for any $x, y \in S$. We call a vector $x$ a convex combination of $x_1, x_2, \ldots, x_m$ if there exists $\lambda_1, \lambda_2, \ldots, \lambda_m \geq 0$ such that $\lambda_1 + \lambda_2 + \cdots + \lambda_m = 1$ and $x = \lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_m x_m$. Obviously, $xy$ is the set consisting of all convex combinations of the two points $x$ and $y$. So $S$ is convex if it contains all convex combinations of any two points in $S$. Actually, this is true for any finitely many points in $S$. That is, we have.

Proposition 2.1. $S$ is convex if and only if it contains all convex combinations of any finitely many points in $S$. The proof of the above proposition...