Linear, affine, and convex sets and hulls

In the sequel, unless otherwise specified, X will denote a real vector space.

Lines and segments. Given two points $x, y \in X$, we define

$$\overleftrightarrow{xy} = \{x + t(y - x) : t \in \mathbb{R}\} = \{(1-t)x + ty : t \in \mathbb{R}\},$$
$$[x, y] = \{x + t(y - x) : t \in [0,1]\} = \{(1-t)x + ty : t \in [0,1]\}.$$

The sets $\overleftrightarrow{xy}$ and $[x, y]$ are, respectively, the *line* passing through x and y, and the (closed) *segment* with endpoints x, y. Observe that they reduce to a singleton whenever $x = y$.

We shall use also the following notation for non-closed segments:

$$(x, y] = [x, y] \setminus \{x\}, \qquad [x, y) = [x, y] \setminus \{y\}, \qquad (x, y) = [x, y] \setminus \{x, y\}.$$

Linear, affine, and convex sets. A set $A \subset X$ is called:

- *linear* if A is a vector subpace of X (i.e., A is nonempty, and $\alpha x + \beta y \in A$ whenever $x, y \in A$, $\alpha, \beta \in \mathbb{R}$);
- *affine* if the line passing through any two points of A is entirely contained in A (i.e., $(1-t)x + ty \in A$ whenever $x, y \in A$, $t \in \mathbb{R}$);
- *convex* if any segment with endpoints in A is contained in A (i.e., $(1-t)x + ty \in A$ whenever $x, y \in A$, $t \in [0,1]$).

Obviously, each linear set is affine, and each affine set is convex. Moreover, any translate of an affine (convex, respectively) set is affine (convex, resp.).

Example 0.1. A linear set in $\mathbb{R}^2$ is either the singleton $\{0\}$, or a line containing 0, or the whole $\mathbb{R}^2$.
An affine set in $\mathbb{R}^2$ is either $\emptyset$, or a singleton, or a line, or $\mathbb{R}^2$.
There is a large variety of convex sets in $\mathbb{R}^2$...

Proposition 0.2. *Let A be a set in a vector space X.*

(a) *A is linear if and only if A is affine and contains* 0.
(b) *A is affine if and only if either $A = \emptyset$ or A is a translate of a linear set. (Moreover, the linear set is unique in this case.)*

Proof. Exercise. ∎

By Proposition 0.2(b), the following definition is justified. The *dimension* and the *codimension* of an affine set $A \subset X$ is defined as, respectively, the dimension and the codimension of the (unique) linear translate of A.

Hyperplanes. A set $H \subset X$ is a *hyperplane* in X if it is an affine set of codimension 1. Equivalently, a hyperplane is any maximal proper affine subset of X.

Proposition 0.3. *A set $H \subset X$ is a hyperplane if and only if it is of the form $H = \varphi^{-1}(\alpha)$ where $\varphi\colon X \to \mathbb{R}$ is a nonzero linear functional and $\alpha \in \mathbb{R}$.*

Proof. "$\Rightarrow$" Fix any $x_0 \in H$. By Proposition 0.2, the set $Y = H - x_0$ is a linear subspace of codimension 1 in X. Fix any $v_0 \in X \setminus Y$. Then each $x \in X$ has a unique representation of the form $x = y_x + t_x v_0$ where $y_x \in Y$, $t_x \in \mathbb{R}$. The mapping $\varphi(x) := t_x$ is linear and satisfies

$$x \in H \Leftrightarrow x - x_0 \in Y \Leftrightarrow \varphi(x - x_0) = 0 \Leftrightarrow \varphi(x) = \varphi(x_0).$$

Thus we can put $\alpha = \varphi(x_0)$.
"$\Leftarrow$" It is easy to see that $\varphi^{-1}(\alpha)$ is a translate of the kernel $\varphi^{-1}(0)$. We have to show that $\varphi^{-1}(0)$ has codimension 1. Fix some $v_0 \in X \setminus \varphi^{-1}(0)$ (this is possible since $\varphi \not\equiv 0$). By substituting v_0 by its appropriate multiple, we can suppose that $\varphi(v_0) = 1$. Then every $x \in X$ can be written in the form

$$x = [x - \varphi(x)v_0] + \varphi(x)v_0 \in \varphi^{-1}(0) + \mathbb{R}v_0$$

since $\varphi(x - \varphi(x)v_0) = 0$. □

Corollary 0.4. *Let H be a hyperplane in X. Then X can be written as a disjoint union*

$$X = H \cup H^+ \cup H^-$$

in such a way that if $x \in H^+$ and $y \in H^-$ then $[x, y] \cap H$ is a singleton. (The sets H^+, H^- are the algebraically open halfspaces generated by *H.)*

Proof. Exercise. Hint: take φ, α as in Proposition 0.3 and consider $H^+ = \{x \in X : \varphi(x) > \alpha\}$, $H^- = \{x \in X : \varphi(x) < \alpha\}$. □

Convex and affine combinations.

Definition 0.5. Let $A \subset X$. An *affine combination of elements of A* is any finite sum of the form

$$\sum_{i=1}^{n} \lambda_i x_i \quad \text{where } x_i \in A,\ \lambda_i \in \mathbb{R},\ \textstyle\sum_1^n \lambda_j = 1. \tag{1}$$

A *convex combination of elements of A* is any finite sum of the form (1) with $\lambda_i \geq 0$ $(i = 1, \dots, n)$.

Proposition 0.6. *Every convex/affine/linear set in a vector space X is closed under making convex/affine/linear combinations of its elements.*

Proof. The "linear part" is well known from Linear Algebra.

Let C be a convex set and $x = \sum_{i=1}^n \lambda_i x_i$ be a convex combination of elements of C. We want to prove that $x \in C$. Let us proceed by induction with respect to n. For $n = 1$, we have $x = x_1 \in C$. Now, suppose that the case $n = k$ holds, and consider the case $n = k + 1$, that is, $x = \sum_{i=1}^{k+1} \lambda_i x_i$ with $x_i \in C$, $\lambda_i \geq 0$,

$\sum_1^{k+1} \lambda_j = 1$. If $\lambda_{k+1} = 1$ then necessarily $x = x_{k+1} \in C$. Suppose $\lambda_{k+1} \neq 1$. Then $s := \sum_1^k \lambda_j = 1 - \lambda_{k+1} \neq 0$. We can write

$$x = (1 - \lambda_{k+1}) \left[\sum_{i=1}^{k} \frac{\lambda_i}{s} x_i \right] + \lambda_{k+1} x_{k+1} \,.$$

Since the sum in the quare brackets belongs to C by our induction assumption, x belongs to C.

The "affine part" can be proved in the same way. The only difference is that we start with indexing the points x_i in such a way that $\lambda_{k+1} \neq 1$ (if this is not possible, we are in a trivial case). □

It is easy to see that the set of all convex combinations of elements of $\{x, y\}$ is the segment $[x, y]$, and the set of all affine combinations of elements of $\{x, y\}$ is the line $\overleftrightarrow{xy}$.

Observation 0.7. *A convex/affine combination of convex/affine combinations of elements of A is a convex/affine combination of elements of A.*

Fact 0.8. *Let X be a normed linear space, $x, y, z \in X$, $z \in (x, y)$. Then*

$$\|x - y\| = \|x - z\| + \|z - y\| \quad \textit{and} \quad z = \tfrac{\|z-y\|}{\|x-y\|} x + \tfrac{\|x-z\|}{\|x-y\|} y \,.$$

Proof. We have $z = (1 - t)x + ty$ for some $t \in (0, 1)$. Then $z - x = t(y - x)$ and $z - y = (1 - t)(x - y)$. Passing to norms we get $t = \frac{\|z-x\|}{\|y-x\|}$ and $1 - t = \frac{\|z-y\|}{\|x-y\|}$. Now the two formulas easily follow. □

Hulls. It is an easy but important observation that the intersection of any family of linear/affine/convex sets is again a linear/affine/convex set. (The same does not hold for unions, but does holds for linearly ordered unions.)

Given a set $A \subset X$, the intersection of all linear sets containing A is the *linear hull* of A, denoted by $\operatorname{span}(A)$. Analogously, we can define the *affine hull* of A as the intersection of all affine sets containing A, and the *convex hull* of A as the intersection of all convex sets containing A. The affine and the convex hull of A will be denoted by

$$\operatorname{aff}(A) \quad \text{and} \quad \operatorname{conv}(A) \,,$$

respectively.

Obviously, A is linear if and only if $\operatorname{span}(A) = A$; A is affine if and only if $\operatorname{aff}(A) = A$; A is convex if and only if $\operatorname{conv}(A) = A$.

It is a well known fact that the linear hull of a set A coincides with the set of all linear combinations of elements of A. The following theorem states that analogous properties hold for convex hulls and for affine hulls as well.

Theorem 0.9. *Let A be a set in a vector space X. Then*

$$\operatorname{conv}(A) = \{x \in X : \ x \text{ is a convex combination of elements of } A\} \,,$$
$$\operatorname{aff}(A) = \{x \in X : \ x \text{ is an affine combination of elements of } A\} \,.$$

Proof. Let us prove the first formula (the second one is analogous). By Observation 0.7, the set C of all convex combinations of points of A is convex; thus $\mathrm{conv}(A) \subset C$. On the other hand, any point $x \in C$, being a convex combination of points of $\mathrm{conv}(A)$, belongs to $\mathrm{conv}(A)$ by Proposition 0.6. $\square$

Let A be an affine set in a vector space X of a finite dimension d and let $x \in \mathrm{aff}(A)$. By translation, we can suppose that $0 \in A$. In this case, x belongs to the linear hull of A, and hence it is a linear combination of at most d elements of A:

$$x = \sum_{i=1}^{d} \lambda_i x_i \quad \text{where } \lambda_i \in \mathbb{R},\ x_i \in A.$$

Since $0 \in A$, we can write x as an affine combination of $d+1$ points of A:

$$x = \lambda_0 \cdot 0 + \sum_{i=1}^{d} \lambda_i x_i \quad \text{with } \lambda_0 = 1 - \textstyle\sum_1^n \lambda_j .$$

Thus we have proved that, *in an d-dimensional vector space, every point of the affine hull of a set is an affine combination of $d+1$ or fewer points of A.* The following important theorem shows that a similar result holds for convex hulls as well.

Theorem 0.10 (Carathéodory)**.** *Let A be a subset of a d-dimensional vector space X. Then*

$$\mathrm{conv}(A) = \{x \in X :\ x \text{ is a convex combination of } d+1 \text{ or fewer points of } A\}.$$

Proof. By Theorem 0.9, it suffices to show that every point of the form

$$x = \sum_{i=0}^{n} \lambda_i x_i \quad \text{where } \lambda_i \in \mathbb{R},\ x_i \in A,\ \sum_{0}^{n} \lambda_j = 1\,,$$

(a convex combination of $n+1$ points of A) is a convex combination of $d+1$ or fewer points of A. If $n \le d$, there is nothing to prove. Let $n > d$. By translation, we can (and do) suppose that $x_0 = 0$. Since the set $\{x_1, \ldots, x_n\}$ is linearly dependent, there exist real numbers $\alpha_1, \ldots, \alpha_n$, not all of them null, such that

$$\sum_{i=1}^{n} \alpha_i x_i = 0\,. \tag{2}$$

Since (2) remains true if we change the sign of all α_i's, we can suppose that $\sum_1^n \alpha_i \ge 0$. Observe that the set $I = \{i \in \{1, \ldots, n\} : \alpha_i > 0\}$ is nonempty. (Indeed, otherwise we would have $\sum_1^n \alpha_i < 0$ since not all a_i's are null.) For each $t > 0$, we have

$$x = \sum_{i=1}^{n} \lambda_i x_i - t \sum_{i=1}^{n} \alpha_i x_i = \sum_{i=1}^{n} (\lambda_i - t\alpha_i) x_i\,.$$

Observe that all coefficients in the last sum will be nonnegative provided $t \le \frac{\lambda_i}{\alpha_i}$ for each $i \in I$. Choose $k \in I$ so that $\frac{\lambda_k}{\alpha_k} = \min\{\frac{\lambda_i}{\alpha_i} : i \in I\}$. Then, for $t = \frac{\lambda_k}{\alpha_k}$, we have

$\lambda_i - t\alpha_i \geq 0$ for each $1 \leq i \leq n$, $\lambda_k - t\alpha_k = 0$, and $\sum_1^n(\lambda_j - t\alpha_j) = \sum_1^n \lambda_j - t\sum_1^n \alpha_j \leq \sum_1^n \lambda_j \leq 1$. Since

$$x = \left[1 - \textstyle\sum_1^n(\lambda_j - t\alpha_j)\right] \cdot 0 + \sum_{\substack{i=1 \\ i\neq k}}^{n} (\lambda_i - t\alpha_i)x_i \,,$$

we have written x as a convex combination of less than $n+1$ points of A.

So, we have proved that, in any convex combination x of more than $d+1$ points, an appropriate change of coefficients allows us to throw out one of the points without changing x. Now, the proof follows by repeating this procedure untill we arrive to at most $d+1$ points. □

Theorem 0.11. *Let X be a normed space of a finite dimension, $K \subset X$ a compact set. Then* $\operatorname{conv}(K)$ *is compact.*

Proof. Let $d = \dim(X)$. Denote $\Lambda = \{\lambda = (\lambda_i)_0^d \in [0,1]^{d+1} : \sum_0^d \lambda_i = 1\}$, and define $F\colon \Lambda \times K^{d+1} \to X$ by $F(\lambda, x_0, \ldots, x_d) = \sum_{i=0}^d \lambda_i x_i$. By the Carathéodory theorem, we have

$$\operatorname{conv}(K) = \{F(\lambda, x_0, \ldots, x_d) : \lambda \in \Lambda,\ x_i \in K\} = F(\Lambda \times K^{d+1}).$$

The last set is compact since F is continuous and $\Lambda \times K^{d+1}$ is a compact metric space. □

Corollary 0.12. *In any normed space, the convex hull of a finite set is compact.* (Indeed, we can restrict ourselves to a finite-dimensional subspace, and apply the above theorem.)

The next example shows that the assumption on the dimension of the space in Theorem 0.11 cannot be omitted.

Example 0.13. Consider the Hilbert space ℓ_2 and a set $K = \{\frac{e_n}{n} : n \in \mathbb{N}\} \cup \{0\}$, where e_n is the n-th vector of the standard orthonormal basis of ℓ_2. The set K is compact since $\frac{e_n}{n} \to 0$. We claim that $\operatorname{conv}(K)$ is not compact since it is not closed. First, the points $x_n := (\sum_1^n 2^{-j})^{-1} \sum_{i=1}^n 2^{-i}(e_i/i)$ $(n \in \mathbb{N})$ belong to K. Second, the sequence $\{x_n\}$ converges in ℓ_2 to the point $x = (x_n)$ with $x_n = 2^{-n}(1/n)$ for each n *(Exercise: prove this!)*. Third, observe that every element of $\operatorname{conv}(K)$ has a finite support; thus $x \notin \operatorname{conv}(K)$.

However, we shall see in a moment that, if the normed space is complete, the closedness is the unique thing which can prevent the convex hull of a compact set from being compact.

Definition 0.14. Let X be a normed space, $A \subset X$. The *closed convex hull* of A is the intersection of all closed convex sets containing A, and it is denoted by $\overline{\operatorname{conv}}(A)$.

Observation 0.15. *Let X be a normed space, $A \subset X$. Then*

$$\overline{\operatorname{conv}}(A) = \overline{\operatorname{conv}(A)}\,.$$

Recall that a metric space (M, d) is *totally bounded* (or *precompact*) if, for each $\varepsilon > 0$, it contains a finite ε-net, that is, a finite set F_ε such that $d(x, F_\varepsilon) < \varepsilon$ for each $x \in M$. It is a well known fact that M is compact if and only if M is complete and totally bounded.

Exercise 0.16. Let (M, d) be a metric space, $A \subset M$.

(a) $\overline{A}$ is totally bounded if and only if A is totally bounded.
(b) A is totally bounded if and only if, for each $\varepsilon > 0$, there exists a compact set $K \subset M$ such that $d(x, K) < \varepsilon$ for each $x \in A$.

Theorem 0.17. *Let X be a normed linear space, $A \subset X$ a totally bounded set.*

(a) $\mathrm{conv}(A)$ *is totally bounded.*
(b) *If X is a Banach space, then* $\overline{\mathrm{conv}}(A)$ *is compact.*

Proof. (a) Fix $\varepsilon > 0$. There exists a finite set $A_0 \subset A$ such that, for each $a \in A$, there exists $y_a \in A_0$ with $\|a - y_a\| < \varepsilon$. The set $\mathrm{conv}(A_0)$ is compact by Corollary 0.12. Now, if $x \in \mathrm{conv}(A)$, we can write $x = \sum_{i=1}^n \lambda_i a_i$ where $a_i \in A$, $\lambda_i \geq 0$, $\sum_1^n \lambda_j = 1$. The point $c = \sum_{i=1}^n \lambda_i y_{a_i}$ belongs to $\mathrm{conv}(A_0)$ and it satisfies

$$\|x - c\| \leq \textstyle\sum_{i=1}^n \lambda_i \|a_i - y_{a_i}\| < \varepsilon \sum_{i=1}^n \lambda_i = \varepsilon\,.$$

By Exercise 0.16(b), $\mathrm{conv}(A)$ is totally bounded.

To show (b) it suffices to observe that $\overline{\mathrm{conv}}(A)$ is complete (since it is closed and X is complete) and totally bounded (by (a) above and Exercise 0.16(a)). □

Now, let us consider convex hulls of finitely many convex sets. We can see the part (b) of Theorem 0.18 as a generalization of the fact that a convex hull of a finite set is compact.

Theorem 0.18. *Let X be a normed space. Let $C_1, \ldots, C_n$ be convex subsets of X.*

(a) $\mathrm{conv}(C_1 \cup \ldots \cup C_n) = \{\sum_{i=1}^n \lambda_i x_i : \ x_i \in C_i, \lambda_i \geq 0, \sum_1^n \lambda_j = 1\}$.
(b) *If each C_i is compact, then* $\mathrm{conv}(C_1 \cup \ldots \cup C_n)$ *is compact.*
(c) *If C_1 is closed and bounded, and the sets $C_2, \ldots, C_n$ are compact, then* $\mathrm{conv}(C_1 \cup \ldots \cup C_n)$ *is closed.*

Proof. (a) The inclusion "$\supset$" is obvious. To see the reverse one, consider $x \in \mathrm{conv}(C_1 \cup \ldots \cup C_n)$ and write it as a convex combination of elements y_k $(k = 1, \ldots, K)$ of $C_1 \cup \ldots \cup C_n$. Since each of y_k's belongs to some C_i, we can group them with respect to which set they belong. Thus x can be written in the form

$$x = \sum_{i=1}^n \left(\sum_{k \in J_i} \lambda_k y_k \right),$$

where J_i's are pairwise disjoint, $\bigcup_1^n J_i = \{1, \ldots, K\}$, $\sum_1^K \lambda_k = 1$, and $x_k \in C_i$ whenever $k \in J_i$.

For every fixed $i \in \{1, \dots, n\}$, denote $\mu_i = \sum_{k \in J_i} \lambda_k$. If $\mu_i = 0$, fix an arbitrary $x_i \in C_i$. If $\mu_i > 0$, denote $x_i = \sum_{k \in J_i} \frac{\lambda_k}{\mu_i} y_k$ and observe that $x_i \in C_i$. Now, we have

$$x = \sum_{i=1}^{n} \mu_i x_i \quad \text{and} \quad \sum_{i=1}^{n} \mu_i = 1 .$$

(b) follows in the same way as in the proof of Theorem 0.11. Indeed, using (a), we can write

$$\operatorname{conv}(C_1 \cup \ldots \cup C_n) = F(\Lambda \times C_1 \times \ldots \times C_n)$$

where $\Lambda = \{\lambda \in [0,1]^n : \sum_1^n \lambda_i = 1\}$ and $F(\lambda, x_1, \dots, x_n) = \sum_{i=1}^n \lambda_i x_i$. Thus $\operatorname{conv}(C_1 \cup \ldots \cup C_n)$ is compact since it is a continuous image of a compact metric space.

(c) Let $\{x_m\} \subset \operatorname{conv}(C_1 \cup \ldots \cup C_n)$ be a sequence converging to some $x \in X$. By (a), each x_m can be written as a convex combination

$$x_m = \sum_{i=1}^{n} \lambda_i^{(m)} c_i^{(m)} \quad \text{where} \quad c_i^{(m)} \in C_i .$$

Using the fact that $[0,1], C_2, \dots, C_n$ are compact, we can pass to subsequences to assure that $\lambda_i^{(m)} \to \lambda_i \in [0,1]$ $(1 \le i \le n)$ and $c_i^{(m)} \to c_i \in C_i$ $(2 \le i \le n)$ as $m \to +\infty$. Observe that $\sum_1^n \lambda_i = 1$. Let us consider two cases.

If $\lambda_1 = 0$, we have $\lambda_1^{(m)} c_1^{(m)} \to 0$ (since C_1 is bounded) and hence

$$x = \lim_m x_m = \lim_m \sum_{i=2}^{n} \lambda_i^{(m)} c_i^{(m)} = \sum_{i=2}^{n} \lambda_i c_i \in \operatorname{conv}(C_2 \cup \ldots \cup C_n) \subset \operatorname{conv}(C_1 \cup \ldots \cup C_n) .$$

If $\lambda_1 > 0$, we can write

$$c_1^{(m)} = \frac{1}{\lambda_1^{(m)}} \left(x_m - \sum_{i=2}^{n} \lambda_i^{(m)} c_i^{(m)} \right) .$$

Thus $c_1^{(m)} \to \frac{1}{\lambda_i} \left(x - \sum_{i=2}^n \lambda_i c_i \right) =: c_1 \in C_1$ (since C_1 is closed). Consequently,

$$x = \lim_m x_m = \sum_{i=1}^{n} \lambda_i c_i \in \operatorname{conv}(C_1 \cup \ldots \cup C_n) .$$

□

Corollary 0.19. *In any normed linear space,* $\operatorname{conv}(C \cup F)$ *is closed whenever* C *is closed, bounded, and convex, and* F *is finite.*

Example 0.20. The assumption that C is bounded in Corollary 0.19 cannot be omitted. To see this, consider a line C in the plane, and a point $x_0 \notin C$. Then $\operatorname{conv}(C \cup \{x_0\})$ is not closed. *(Why?)*