

Lauritzen: *Undergraduate Convexity*

Danny Nygård Hansen

16th December 2023

2 • Affine subspaces

General notes

REMARK 2.1. Let $f: \mathbb{R}^d \rightarrow \mathbb{R}^m$ be a map. Then the following properties are equivalent:

- (a) There exists a matrix A and a vector b such that $f(x) = Ax + b$ for all $x \in \mathbb{R}^d$.
- (b) For all $x, y \in \mathbb{R}^d$ and $t \in \mathbb{R}$, $f((1-t)x + ty) = (1-t)f(x) + tf(y)$.
- (c) For all $x_1, \dots, x_k \in \mathbb{R}^d$ and $t_1, \dots, t_k \in \mathbb{R}$ with $t_1 + \dots + t_k = 1$, $f(t_1x_1 + \dots + t_kx_k) = t_1f(x_1) + \dots + t_kf(x_k)$.

Such a function is called *affine*. Note that if $m = 1$, then f is affine if and only if it is both convex and concave.

The first property clearly entails the third, which then entails the second, so assume that f has the second property. First assume that $f(0) = 0$. For $\beta \in \mathbb{R}$ and $x \in \mathbb{R}^d$ we thus have

$$f(\beta x) = f(\beta x + (1-\beta)0) = \beta f(x) + (1-\beta)f(0) = \beta f(x),$$

so f is homogeneous. If also $y \in \mathbb{R}^d$, then

$$\frac{1}{2}f(x+y) = f\left(\frac{1}{2}x + \frac{1}{2}y\right) = \frac{1}{2}f(x) + \frac{1}{2}f(y),$$

so f is also additive, hence linear. For general f , simply replace f with $f - f(0)$. ┘

REMARK 2.2. Let $f: \mathbb{R}^d \rightarrow \mathbb{R}^m$ be affine, and let $M \subseteq \mathbb{R}^d$ be affine. Then the image $f[M]$ is also affine. Furthermore, for any $S \subseteq \mathbb{R}^d$ we have $f[\text{aff } S] = \text{aff } f[S]$ (just as linear maps commute with span).

If $f(x) = Ax + b$ and $M = x_0 + W$ is affine with W a subspace, then

$$f[M] = f[x_0 + W] = A[x_0 + W] + b = (Ax_0 + b) + A[W].$$

In particular, $\dim f[M] = \dim A[W]$. If f (and hence A) is invertible, then we thus have $\dim f[M] = \dim W = \dim M$. \lrcorner

REMARK 2.3. If $S \subseteq \mathbb{R}^d$ and $x_0 \in S$, then we claim that $\text{aff } S = x_0 + \text{span}(S - x_0)$. Consider the affine map $f(x) = x - x_0$. From the above remark we have

$$x_0 + \text{aff } f[S] = x_0 + f[\text{aff } S] = \text{aff } S.$$

But $f[S]$ contains 0, so $\text{aff } f[S]$ is a subspace and hence equal to $\text{span } f[S]$. Furthermore, since f is invertible it preserves dimensions, so if $|S| = m$ and is positive and finite, then $\dim S = m - 1$ if and only if $\dim W = m - 1$, where $W = \text{span}(S - x_0)$. But W is also spanned by $S \setminus \{x_0\} - x_0$ which has $m - 1$ elements, so S is affinely independent if and only if $S \setminus \{x_0\} - x_0$ is linearly independent. \lrcorner

REMARK 2.4. If $f: \mathbb{R}^d \rightarrow \mathbb{R}^d$ preserves the Euclidean metric, then it is affine. Replacing f with $f - f(0)$ we may assume that $f(0) = 0$, and so that f also preserves the inner product. For all $x, y, z \in \mathbb{R}^d$ and $\beta \in \mathbb{R}$ we have

$$\begin{aligned} \langle f(\beta x + y), f(z) \rangle &= \langle \beta x + y, z \rangle \\ &= \beta \langle x, z \rangle + \langle y, z \rangle \\ &= \beta \langle f(x), f(z) \rangle + \langle f(y), f(z) \rangle \\ &= \langle \beta f(x) + f(y), f(z) \rangle. \end{aligned}$$

Since f is surjective, the claim follows. \lrcorner

REMARK 2.5: The affine hull and dimension of the empty set.

By definition, the affine hull of the empty set is the empty set. This is the case by Definition 2.4, and this is also the standard definition. Hence the affine hull of the empty set is not an affine subspace! In particular, it cannot have a dimension in the sense of Definition 2.7. \lrcorner

Exercises

EXERCISE 2.4

Prove that $M = \{v \in \mathbb{R}^d \mid Av = b\}$ is an affine subspace of \mathbb{R}^d , where A is an $m \times d$ matrix and $b \in \mathbb{R}^m$.

SOLUTION. For $u, v \in \mathbb{R}^d$ and $t \in \mathbb{R}$ we have

$$A((1-t)u + tv) = (1-t)Au + tAv = (1-t)b + tb = b. \quad \square$$

EXERCISE 2.10

Prove that you can have no more than $d + 1$ affinely independent vectors in \mathbb{R}^d .

SOLUTION. Let $\{v_1, \dots, v_n\} \subseteq \mathbb{R}^d$ be affinely independent. By Proposition 2.9, the $n - 1$ vectors $v_2 - v_1, \dots, v_n - v_1$ are linearly independent. But then we must have $n - 1 \leq d$, i.e. $n \leq d + 1$. \square

EXERCISE 2.11

Let v_0, \dots, v_d be affinely independent points in \mathbb{R}^d . Prove that

$$f(x) = (\lambda_0, \lambda_1, \dots, \lambda_d)$$

is a well defined affine map $f: \mathbb{R}^d \rightarrow \mathbb{R}^{d+1}$, where

$$x = \lambda_0 v_0 + \dots + \lambda_d v_d$$

with $\lambda_0 + \dots + \lambda_d = 1$.

SOLUTION. If also $x = \mu_0 v_0 + \dots + \mu_d v_d$ with $\mu_0 + \dots + \mu_d = 1$, then

$$(\lambda_0 - \mu_0)v_0 + \dots + (\lambda_d - \mu_d)v_d = 0,$$

so since the v_i are affinely independent we have $\lambda_i = \mu_i$ by Proposition 2.9. Hence f is well-defined.

Next we show that f is affine. Let

$$x = \lambda_0 v_0 + \dots + \lambda_d v_d \quad \text{and} \quad y = \mu_0 v_0 + \dots + \mu_d v_d$$

with $\lambda_0 + \dots + \lambda_d = 1$ and $\mu_0 + \dots + \mu_d = 1$, and let $t \in \mathbb{R}$. Then

$$\sum_{i=0}^d ((1-t)\lambda_i + t\mu_i) = (1-t) \sum_{i=0}^d \lambda_i + t \sum_{i=0}^d \mu_i = (1-t) + t = 1.$$

Therefore, since

$$(1-t)x + ty = \sum_{i=0}^d ((1-t)\lambda_i + t\mu_i)v_i,$$

we have

$$\begin{aligned} f((1-t)x + ty) &= ((1-t)\lambda_0 + t\mu_0, \dots, (1-t)\lambda_d + t\mu_d) \\ &= (1-t)(\lambda_0, \dots, \lambda_d) + t(\mu_0, \dots, \mu_d) \\ &= (1-t)f(x) + tf(y), \end{aligned}$$

so f is affine. \square

EXERCISE 2.12

Prove that a non-empty open subset $U \subseteq \mathbb{R}^d$ has dimension $\dim U = d$. Show that a subset $S \subseteq \mathbb{R}^d$ with $\dim S = d$ contains a non-empty open subset.

SOLUTION. Let $x \in U$, and let (e_1, \dots, e_d) be the standard basis for \mathbb{R}^d . Since U is open there is some $\delta > 0$ such that $x + \delta e_i \in U$ for all i . The linear subspace $-x + \text{aff } U$ has dimension d , so U has dimension d .

The latter claim is in fact false, since if S is e.g. the collection of points δe_i along with the origin, then S has dimension d (since $\text{aff } S$ contains each of the coordinate axes, hence the entire \mathbb{R}^d) but clearly contains no non-empty open subset. \square

3 • Convex subsets

Exercises

EXERCISE 3.7

We state and prove a more general result than the one given in the exercise.

For any subsets $A, B \subseteq \mathbb{R}^d$ we have

$$\text{conv}(A + B) = \text{conv } A + \text{conv } B.$$

SOLUTION. The inclusion ' \subseteq ' is clear since the sum of convex sets is convex. For the converse direction, let $v \in \text{conv } A$ and $w \in \text{conv } B$, and write $v = \sum_{i=1}^n \lambda_i v_i$ and $w = \sum_{j=1}^m \mu_j w_j$ as convex combinations of elements in A and B respectively. Then

$$v + w_j = v + (\lambda_1 + \dots + \lambda_n)w_j = \sum_{i=1}^n \lambda_i(v_i + w_j) \in \text{conv}(A + B),$$

and

$$v + w = (\mu_1 + \cdots + \mu_n)v + w = \sum_{j=1}^m \mu_j(v + w_j) \in \text{conv conv}(A + B) = \text{conv}(A + B),$$

as desired. \square

EXERCISE 3.9

We state and prove a more general result than the one given in the exercise.

For $S \subseteq \mathbb{R}^d$, prove that $(\text{conv } S)^\circ = S^\circ$. If $S = \{v_1, \dots, v_n\}$ is finite, then

$$(\text{conv } S)^\circ = \{\alpha \in \mathbb{R}^d \mid \alpha^\top v_1 \leq 1, \dots, \alpha^\top v_n \leq 1\} = \{\alpha \in \mathbb{R}^d \mid A\alpha \leq 1\},$$

where A is the $n \times d$ matrix with rows $v_1^\top, \dots, v_n^\top$. In particular, polytopes are polyhedra.

SOLUTION. Since $S \subseteq \text{conv } S$, we clearly have $(\text{conv } S)^\circ \subseteq S^\circ$. For the opposite inclusion, let $\alpha \in S^\circ$ and let $\lambda_1 v_1 + \cdots + \lambda_n v_n$ be a convex combination of elements in S . Then

$$\alpha^\top (\lambda_1 v_1 + \cdots + \lambda_n v_n) = \lambda_1 \alpha^\top v_1 + \cdots + \lambda_n \alpha^\top v_n \leq \lambda_1 + \cdots + \lambda_n = 1,$$

so $\alpha \in (\text{conv } S)^\circ$. The rest of the claims follow easily. \square

EXERCISE 3.10

If $F \subseteq G \subseteq C$ are convex subsets of \mathbb{R}^d , prove that F is a face of C if F is a face of G and G is a face of C .

SOLUTION. Let $x, y \in C$ such that the open line segment (x, y) intersects F . Then it intersects G , so since G is a face of C we have $x, y \in G$. Similarly, since F is a face of G we then have $x, y \in F$. Hence F is a face of C . \square

EXERCISE 3.12

Prove that $C \setminus F$ is a convex subset if F is a face of a convex subset C . Is it true that $F \subseteq C$ is a face if $C \setminus F$ is a convex subset?

SOLUTION. Let $x, y \in C \setminus F$. If the open line segment (x, y) intersected F , then we would have $x, y \in F$ since F is a face of C . Thus $(x, y) \subseteq C \setminus F$, so in total $[x, y] \subseteq C \setminus F$. Hence $C \setminus F$ is convex.

The converse is false. For instance, remove (at least) two extreme points from a disk. \square