6 Affine Sets II

Definition. If $\mathbf{x_1}, \ldots, \mathbf{x_k} \in \mathbb{R}^n$ then an *affine combination* of $\mathbf{x_1}, \ldots, \mathbf{x_k}$ is a linear combination

$$c_1\mathbf{x_1} + \ldots + c_k\mathbf{x_k}$$

with the additional property that $c_1 + \ldots + c_k = 1$. The *affine hull* of $\mathbf{x_1}, \ldots, \mathbf{x_k}$ is the set of all affine combinations of these points, denoted

$$\mathrm{Aff}(\mathbf{x_1}, \ldots, \mathbf{x_k}) = \{c_1\mathbf{x_1} + \ldots + c_k\mathbf{x_k} \mid c_1 + \ldots + c_k = 1\}.$$

Examples:

1. For a single point $\mathbf{x_1} \in \mathbb{R}^n$ the definition gives $\mathrm{Aff}(\mathbf{x_1}) = \{c_1\mathbf{x_1} \mid c_1 = 1\} = \{\mathbf{x_1}\}$.

2. If $\mathbf{x_1}, \mathbf{x_2} \in \mathbb{R}^n$ are distinct, the set $\mathrm{Aff}(\mathbf{x_1}, \mathbf{x_2})$ is a set we have already met—it is the line through $\mathbf{x_1}$ and $\mathbf{x_2}$. This fact follows from the equation

$$\begin{aligned}
\mathrm{Aff}(\mathbf{x_1}, \mathbf{x_2}) &= \{c_1\mathbf{x_1} + c_2\mathbf{x_2} \mid c_1 + c_2 = 1\} \\
&= \{\mathbf{x_1} + (c_1 - 1)\mathbf{x_1} + c_2\mathbf{x_2} \mid c_1 + c_2 = 1\} \\
&= \{\mathbf{x_1} - c_2\mathbf{x_1} + c_2\mathbf{x_2} \mid c_2 \in \mathbb{R}\} \\
&= \{\mathbf{x_1} + c_2(\mathbf{x_2} - \mathbf{x_1}) \mid c_2 \in \mathbb{R}\} \\
&= \mathbf{x_1} + \{c_2(\mathbf{x_2} - \mathbf{x_1}) \mid c_2 \in \mathbb{R}\} \\
&= \mathbf{x_1} + \mathrm{Span}(\mathbf{x_2} - \mathbf{x_1})
\end{aligned}$$

3. For three distinct points $\mathbf{x_1}, \mathbf{x_2}, \mathbf{x_3} \in \mathbb{R}^n$ we will prove that the set $\mathrm{Aff}(\mathbf{x_1}, \mathbf{x_2}, \mathbf{x_3})$ will be a line if all three of these points lie on a common line. Otherwise $\mathrm{Aff}(\mathbf{x_1}, \mathbf{x_2}, \mathbf{x_3})$ will be a plane.

4. In general we will show that $\mathrm{Aff}(\mathbf{x_1}, \ldots, \mathbf{x_k})$ is the smallest affine set containing the points $\mathbf{x_1}, \ldots, \mathbf{x_k}$.

Lemma 6.1. *The affine hull of* $\mathbf{x_1}, \ldots, \mathbf{x_k}$ *is affine.*

Proof. Let $\mathbf{y}, \mathbf{z} \in \text{Aff}(\mathbf{x_1}, \ldots, \mathbf{x_k})$ be distinct and let $\mathbf{w}$ be a point on the line $\overleftrightarrow{\mathbf{yz}}$. Since $\mathbf{w}$ is on this line we may choose a real number t so that $\mathbf{w} = t\mathbf{y} + (1-t)\mathbf{z}$. Since $\mathbf{y}$ and $\mathbf{z}$ lie in the affine hull we may choose $c_1, \ldots, c_k, d_1, \ldots, d_k \in \mathbb{R}$ so that

$$\mathbf{y} = c_1\mathbf{x_1} + \ldots + c_k\mathbf{x_k}$$
$$\mathbf{z} = d_1\mathbf{y_1} + \ldots + d_k\mathbf{x_k}$$

with the added property $c_1 + \ldots + c_k = 1$ and $d_1 + \ldots + d_k = 1$. Now we have

$$\begin{aligned}\mathbf{w} &= t\mathbf{y} + (1-t)\mathbf{z}\\ &= t\Big(c_1\mathbf{x_1} + \ldots + c_k\mathbf{x_k}\Big) + (1-t)\Big(d_1\mathbf{y_1} + \ldots + d_k\mathbf{x_k}\Big)\\ &= (tc_1 + (1-t)d_1)\mathbf{x_1} + \ldots + (tc_k(1-t)d_k)\mathbf{x_k}\end{aligned}$$

The sum of the coefficients in the above linear combination is

$$(tc_1 + (1-t)d_1) + \ldots + (tc_k(1-t)d_k) = t(c_1 + \ldots + c_k) + (1-t)(d_1 + \ldots + d_k) = t + (1-t) = 1$$

so we find that $\mathbf{w} \in \text{Aff}(\mathbf{x_1}, \ldots, \mathbf{x_k})$ as desired. □

Theorem 6.2. *The affine hull of* $\mathbf{x_1}, \ldots, \mathbf{x_k}$ *is the unique minimal affine set containing these points.*

Proof. To prove this theorem, it suffices to show that every affine set U containing $\mathbf{x_1}, \ldots, \mathbf{x_k}$ also contains $\text{Aff}(\mathbf{x_1}, \ldots, \mathbf{x_k})$. We do this by induction on k.

□

Theorem 6.3. *Every affine set can be expressed as an intersection of hyperplanes.*

Proof. Let U be an affine set in $\mathbb{R}^n$. If $U = \emptyset$ then we can write U as an intersection of two disjoint hyperplanes (we assume here that $n \geq 1$). Otherwise, we have $U = \mathbf{w} + V$ where V is a subspace. It follows from a result in linear algebra that there exists a matrix A so that V is precisely the Nullspace of A.[1] Define the vector $\mathbf{y}$ by the equation

$$A\mathbf{w} = \mathbf{y}.$$

[1]To prove this property, choose an orthogonal basis $\{\mathbf{v_1}, \ldots, \mathbf{v_k}\}$ of V and extend this to an orthogonal basis $\{\mathbf{v_1}, \ldots, \mathbf{v_n}\}$ of $\mathbb{R}^n$. Now the matrix A can be constructed by taking the vectors $\mathbf{v_{k+1}}, \ldots, \mathbf{v_n}$ as rows.

Claim: $U = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \mathbf{y}\}$

The proof of this claim follows immediately from the following equivalence.

$$\begin{aligned} A\mathbf{x} = \mathbf{y} &\Leftrightarrow A\mathbf{x} - A\mathbf{w} = \mathbf{0} \\ &\Leftrightarrow A\left(\mathbf{x} - \mathbf{w}\right) = \mathbf{0} \\ &\Leftrightarrow \mathbf{x} - \mathbf{w} \in V \\ &\Leftrightarrow \mathbf{x} \in \mathbf{w} + V \\ &\Leftrightarrow \mathbf{x} \in U \end{aligned}$$

Next express A (in terms of row vectors) and $\mathbf{y}$ as follows

$$A = \begin{bmatrix} \mathbf{a_1} \\ \vdots \\ \mathbf{a_m} \end{bmatrix} \qquad \mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}$$

Now applying the claim gives us

$$\begin{aligned} U &= \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \mathbf{y}\} \\ &= \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a_i} \cdot \mathbf{x} = y_i \text{ holds for every } 1 \leq i \leq m\} \\ &= \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} \in H_{\mathbf{a_i}}^{y_i} \text{ holds for every } 1 \leq i \leq m\} \\ &= H_{\mathbf{a_1}}^{y_1} \cap \ldots \cap H_{\mathbf{a_m}}^{y_m} \end{aligned}$$

and this completes the proof. □