Let S be a set of points.

If we consider the set $AffineHull_2(S) = \{x \in X : x = \lambda x_1 + (1-\lambda)x_2, \forall x_1, x_2 \in S, \lambda \in \mathcal{R}\}$ to be the affine hull, can we prove or disprove that $AffineHull_2(AffineHull_2(S)) = AffineHull_2(S)$?

Example: the data augmentation technique, Mixup

If not, how do if we can achieve the affine hull in a finite number of steps? Is it a problem that the affine hull does not use a countable/uncountable affine combination of points?

What if S is not a closed set?

Now let C be an affine set. Let's try to understand what's going on in Boyd's proof that affine sets are linear subspaces with an offset.

First, we define what V is.

We can define the translation $\tau_{x_0}(v) = v - x_0$, and we can extend the notation to $\tau_{x_0}(C) = \{\tau_{x_0} v, \forall v \in C\}$.

Notably, τ_{x_0} is a one-to-one mapping. Essentially, Boyd's proof is mapping points from C to V and proving statements in the transformed space. If it were not a one to one mapping, it would be hard to prove things this way, since points in C could have multiple or no corresponding points in V and vice-versa.

Boyd tries to prove that $V = \tau_{x_0}(C)$ is a linear subspace. To prove something is a linear subspace, we have to verify 3 properties:

- $\exists 0 \in V$
- closure under homogeneity (scalar multiplication)
- closure under addition (additivity)

We can accomplish all of these by proving $\forall \alpha, \beta \in R, \forall v_1, v_2 \in V, \alpha v_1 + \beta v_2 \in V$. (Quick question, how does this prove the existence of a 0 element?)?

Now we know that v_1, v_2 have corresponding elements in C, specifically, $\tau_{x_0}^{-1}(v_1) = v_1 + x_0, \tau_{x_0}^{-1}(v_2) = v_2 + x_0$

We want to prove that $\tau_{x_0}^{-1}(\alpha v_1 + \beta v_2) \in C$.

First, let's try a more direct approach: is $\tau_{x_0}^{-1}$ a linear transformation?

That is, is $\tau_{x_0}^{-1}(\alpha v_1 + \beta v_2) = \alpha \tau_{x_0}^{-1}(v_1) + \beta \tau_{x_0}^{-1}(v_2)$?

The answer is no:

$$\tau_{x_0}^{-1}(\alpha v_1 + \beta v_2) = \alpha v_1 + \beta v_2 + x_0 \neq \alpha(v_1 + x_0) + \beta(v_2 + x_0) \tag{1}$$

But our $v_1, v_2 \notin C$, so we have to map them into and replace them with corresponding points in C.

$$\begin{aligned} \alpha \tau_{x_0}^{-1}(v_1) + \beta \tau_{x_0}^{-1}(v_2) &= \alpha v_1 + \beta v_2 + (\alpha + \beta)x_0 \\ \tau_{x_0}^{-1}(\alpha v_1 + \beta v_2) &= \alpha \tau_{x_0}^{-1}(v_1) + \beta \tau_{x_0}^{-1}(v_2) - (\alpha + \beta)x_0 + x_0 \\ &= \alpha \tau_{x_0}^{-1}(v_1) + \beta \tau_{x_0}^{-1}(v_2) + (1 - \alpha - \beta)x_0 \end{aligned} \tag{2}$$

And we see that it is an affine combination of three points. Notably this proof would work if $0 < \alpha, \beta$ and $\alpha + \beta < 1$, that is for convex combinations as well.

Note that we could not directly prove V was linear in its space, we had to do it C space.

Another exercise, let's try to directly prove the converse, if V is linear, then $\tau_{x_0}^{-1}(V) = C$ is affine:

This seems easier, because we need to prove that affine combinations of points in $\tau_{x_0}^{-1}(V)$ are in C. Mathematically, we do this by moving points in V space to C, projecting the result back to V and checking to see if it's in V:

$$\tau_{x_0}\left(\lambda\tau_{x_0}^{-1}(v_1) + (1-\lambda)\tau_{x_0}^{-1}(v_2)\right) \in V \tag{3}$$

Let's begin with the term inside the parentheses:

$$\begin{aligned} &\lambda\tau_{x_0}^{-1}(v_1) + (1-\lambda)\tau_{x_0}^{-1}(v_2) \\ &= \lambda(v_1 + x_0) + (1-\lambda)(v_2 + x_0) \\ &= \lambda v_1 + (1-\lambda)v_2 + x_0 \end{aligned} \tag{4}$$

And translating that by x_0 to V space, we get $\tau_{x_0}\left(\lambda v_1 + (1-\lambda)v_2 + x_0\right) = \lambda v_1 + (1-\lambda)v_2$, which by linearity is in V.

Note the difference, checking that V is linear requires affine combinations while checking that C is affine requires affine combinations, which is follows directly from linearity.

My next question, if we start with a linear combination in V, can we convert it into an affine combination?

More formally, we start with scalar coefficients $\alpha, \beta \in R$, $v_1, v_2 \in V$, and our offset $x_0 \in C$. How can we find the corresponding affine combination in C?

Technically we already have our answer:

$$\tau_{x_0}^{-1}(\alpha v_1 + \beta v_2) = \alpha\tau_{x_0}^{-1}(v_1) + \beta\tau_{x_0}^{-1}(v_2) + (1-\alpha-\beta)x_0 \tag{5}$$

But what exactly does this mean? It means we are augmentation our linear combination with a 3rd term:

$$\alpha v_1 + \beta v_2 + (1-\alpha-\beta)0 \tag{6}$$

and the 0 is mapping to x_0. And this brings up a trivial but pretty interesting point:

we can place a multiple αx_0 into C by taking another point $c_1 \in C$ and adding it with an appropriate scalar multiple:

$$\alpha x_0 + (1-\alpha)c_1 \in C \tag{7}$$

That is, we can't add arbitrary multiples of our offset x_0 and stay with our affine set C, but if we mix it appropriately, we can, using the affine property.

Next, how do we know that for different $x_0, x_1 \in C$, $\tau_{x_0}(C) = \tau_{x_1}(C)$?

Prove the converse: for every affine set, we can express it as the solution set to a system of linear equations, $\{x \in X : Ax = b\}$. Specifically, how do we find A and b?

And what does this have to do with offsets?