As we “safely” leave $\mathbb{R}$ for higher dimensions, we wish to extend many of the concepts we’ve used. Unfortunately, these concepts become more complex. For many applications we will need a new notion of distance.

In particular, we will focus on extending our definition of a median. This is not as intuitive of an extension as with other topics (e.g. nearest neighbor, Delaunay triangulations, Voronoi diagrams), but we will use the property that in one-dimension we based our medians on percentiles.

So, given $P = \{p_1, \ldots, p_n\} \subset \mathbb{R}^d$, we want to find some sort of median of these points. However, it’s hard to determine who is “bigger” when comparing two points in any dimension greater than 1. In order to get around this, we’ll compare points after projecting them onto a one-dimensional subspace.

For our first attempt, we will investigate such a projection of points in $\mathbb{R}^2$. Note that our representation of a point $(x, y) \in \mathbb{R}^2$ is a pair of projects (onto the $x$ and $y$-axes). If we look only at the $x$-coordinates, we have reduced our frame of view to a one-dimensional space. This is true for projections onto the span of one vector in any direction. Using each direction possible, we generate an $\infty$-collection of constraints. For each of these constraints, we have a region in which the projected space where a median could lie (from our 1-dimensional construction of a median), the pre-image of which is a stripe in our 2-dimensional space. Thus, our collection of constraints gives us an $\infty$-collection of stripes from which our median-points could lie in.

More generally, in an $n$-dimensional space we define half-spaces by specifying the hyperplane that cuts the space. For example, in 1-d, the half spaces are defined by a point. Likewise, any 2-d space is split into half spaces by a line, and a plane in 3-space generates two half spaces.

How can we use these to define a median? If we can define the notion of a median in $n-1$-dimensions, then we can project our $n$-dimensional space into the hyperplane, and use the $n-1$-dimension rule to determine a median.

A hyperplane normal to some vector $v$ defines two half spaces: $H^+, H^-$, where $H^+$ consists of all vectors on the same side of the hyperplane as $v$. From this, we will describe our generalization of a median, referred to as a centerpoint.

$c$ is a $\delta$-centerpoint if $\forall$ hyperplanes $H$ containing $c$, $\frac{|H^+ \cap P|}{|P|} \leq 1 - \delta$ (see Figure 1). This is equivalent to saying that $c$ has to be a $\delta$-centerpoint with respect to the projection onto any hyperplane.

We can see the following pattern (as in Figure 2). In 1-dimension, there is always a $\frac{1}{2}$-median. In 2-d, there is always a $\frac{1}{3}$-centerpoint. In 3-d there must exist a $\frac{1}{4}$-centerpoint. Is it true that in $d$-dimensions there must be a $\frac{1}{d+1}$-centerpoint?

**Theorem:** $\forall P \subset \mathbb{R}^{d+1}$, $\exists$ a $\frac{1}{d+1}$-centerpoint.

**Proof:** We will use a recursive Divide-and-Conquer method which results in an algorithm to find a point. The first tool we need is a result from convex-
geometry: Helly’s Theorem.

Helly’s Theorem: Let $C = \{C_1, C_2, \ldots, C_n\}$ such that $C_i$ is a convex set in $\mathbb{R}^d$. If, $\forall(d + 1)$-tuples $(i_1, i_2, \ldots, i_{d+1})$, $\bigcap_{j=1}^{n} C_{i_j} \neq \emptyset$, then $\bigcap_{i=1}^{n} C_i \neq \emptyset$.

Thus, if there is always a “$d + 1$-wise” intersection, then the sets must all intersect.

We will wait to prove for a bit. However, we will look at a few examples in small dimensions. In 1-d the convex sets are simply intervals, and if every pair intersect, then the whole collection intersects. However, in 2-d, we can easily find intervals which pairwise intersect, but do not totally intersect (figure 1). (As soon as we enforce that they must 3-wise intersect, we can’t find such an example.)

We will show a quick 1-dimensional proof. Let $I_1, I_2, I_3$ be intervals in $\mathbb{R}$. Since their pairwise intersections are non-empty, there must be $a_3 \in I_1 \cap I_2, a_2 \in I_1 \cap I_3, a_1 \in I_2 \cap I_3$. Let $a_k = \text{median}(a_1, a_2, a_3)$. Without loss of generality, let
$k = 2.$

\[ a_1 \in I_2 \quad a_2 \quad a_3 \in I_2 \]

Figure 3: Since $I_2$ is convex, $a_2$ must lie within it.

Now, since $a_1$ and $a_3$ are in $I_2$, and $I_2$ is convex, it must be that $a_2 \in I_2$. So $a_2 \in I_1 \cap I_2 \cap I_3$ (see Figure 3).

We will use Helly’s Theorem to show the existence of our $\frac{1}{d+1}$-centerpoint.

Intuitively, $c$ is a $\delta$-centerpoint $\Rightarrow \forall$ half space $H^+: \text{if } \frac{|H^+ \cap P|}{|P|} > 1 - \delta$ then $c \in H^+$, meaning “We’ve made a good split.”

Let $H = \{h_1, \ldots, h_\infty\}$ be a collection of all $\frac{1}{d}$-spaces so $\frac{|h_i \cap P|}{|P|} \geq \frac{1}{d+1}$. We want to show that $\bigcap_{i=1}^\infty h_i \neq \emptyset$. If this is true, then $\exists c \in \mathbb{R}^d$ such that $\forall i : c \in h_i$, and $c$ can be our centerpoint. We will do this via Helly’s Theorem.

Thus, it remains to show that $\forall H = \{h_{i_1}, \ldots, h_{i_{d+1}}\} \subset H, \exists c \in H$, or simply $h_{i_1} \cap h_{i_2} \cap \cdots \cap h_{i_{d+1}} \neq \emptyset$. If we look at the complements of these sets, then we want to show that $(\mathbb{R}^d \setminus h_{i_1}) \cup (\mathbb{R}^d \setminus h_{i_2}) \cup \cdots \cup (\mathbb{R}^d \setminus h_{i_{d+1}}) \neq \mathbb{R}^d$.

We can do this by considering the set $P \cap \bigcup_{i=1}^\infty \overline{h_i} = \bigcup (\overline{h_i} \cap P)$. From above, we know that $\forall i : |h_i \cap P| \geq \frac{d}{d+1} |P|$, or $|\overline{h_i} \cap P| < \frac{1}{d+1} |P|$. Thus, $\bigg| \bigcup_{i=1}^\infty h_i \cap P \bigg| < |h_{i_1} \cap P| + \cdots + |h_{i_{d+1}} \cap P| < (d+1) \frac{1}{d+1} |P|$.

Thus, there must be some $p \in P$ such that $p \in \bigcap h_i$, and we can use Helly’s Theorem!

For next time, we will build an algorithm from this proof.