Why is convexity useful? Linear Separability!

We define convex sets, functions, and problems. One benefit of convexity is that it enables binary search.

A set K in $\mathbb{R}^n$ is *convex* if for every pair of points $x, y \in K$, we have $[x, y] \subseteq K$.

A function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is *convex* if for any $t \in [0, 1]$, we have

$$f(t \cdot x + (1 - t) \cdot y) \leq t \cdot f(x) + (1 - t) \cdot f(y)$$

An optimization problem $\min\limits_{x \in K} f(x)$ is convex if K and f are convex.

One property of convexity states that for any point not in the convex set, we can define a separating hyperplane such that the set is on one side of the hyperplane, and the point is on another. We will use this fact in the future to show that this property allows us to do binary search to find a point in a convex set. Geometrically, we can think of a convex set as a polygon. If we have some point that is not in the polygon, then we can draw a line to partition the space into two distict partitions, one of which contains the point and the other of which contains the convex set in its entirety.

Hyperplane Separation Theorem
Let K be a closed convex set in $\mathbb{R}^n$ and $y \notin K$. Then there exists some non-zero $\theta \in \mathbb{R}^n$ such that $\langle 0, y \rangle > \max\limits_{x \in K} \langle \theta, x \rangle$.

///// Geometric Intuition /////

To gain some intuition, let K be a nonempty closed convex set in $\mathbb{R}^2$, and let y be a point not in K. Since y is not in K, by the Hyperplane Separation Theorem, there exists a hyperplane that

separates y from K.

To find this hyperplane, we take the vector $\theta = y - x$, where x is an arbitrary point in K. Then, we set the inequality $\langle \theta, y \rangle > \max_{x \in K} \langle \theta, x \rangle$, which defines the hyperplane. This hyperplane separates y from K because for any point $x \in K$, we have $\langle \theta, y \rangle > \langle \theta, x \rangle$.

///// Proof /////

The Hyperplane Separation Theorem states that for any nonempty closed convex set K in $\mathbb{R}^n$ and any point y not in K, there exists a non-zero vector $\theta \in \mathbb{R}^n$ such that $\langle \theta, y \rangle > \max_{x \in K} \langle \theta, x \rangle$. The proof of this theorem relies on the fact that for any closed convex set, there exists a point in the set that is closest to a given point outside the set. This point is known as the projection of the given point onto the set, and it is unique for closed convex sets.

To prove the Hyperplane Separation Theorem, we first show the existence of the projection of a point y onto a closed convex set K. We can use the fact that the set of all projections onto K of a point y, denoted by $P_K(y)$, is the set of all points $x \in K$ that minimize the distance between y and x. Formally,

$$P_K(y) = \{x \in K \,|\, \|y - x\|_2 = \min_{z \in K} \|y - z\|_2\}$$

By the properties of convex sets, the function $f(x) = \|y - x\|_2^2$ is a convex function, and thus its global minimum exists and is unique. Therefore, the set of all projections of a point onto a closed convex set is non-empty, and it contains exactly one point, which we denote by x^*. This point x^* is the projection of y onto the closed convex set K.

We can now use the existence of the projection of a point onto a closed convex set to prove the Hyperplane Separation Theorem. Let K be a nonempty closed convex set in $\mathbb{R}^n$, and let y be a point not in K. We define the vector $\theta = y - x^*$, where x^* is the projection of y onto K. Using the properties of the projection, we have that for all

$x \in K$, the dot product of θ and the vector $y - x$ is non-negative:

$$\langle \theta, y - x \rangle = \langle y - x^*, y - x \rangle = \|y - x^*\|_2^2 - \|x^* - x\|_2^2 \geq 0$$

As y is not in K, the vector θ is non-zero. Therefore, we can divide both sides of the above inequality by $\|\theta\|^2$, which is positive, to get:

$$\frac{\langle \theta, y - x \rangle}{\|\theta\|^2} = \frac{\|y - x^*\|_2^2 - \|x^* - x\|_2^2}{\|y - x^*\|_2^2} = 1 - \frac{\|x^* - x\|_2^2}{\|y - x^*\|_2^2} \geq 0$$

This implies that

$$\langle \theta, y \rangle = \langle \theta, y - x^* + x^* \rangle$$

$$= \langle \theta, y - x^* \rangle + \langle \theta, x^* \rangle$$

$$= \|\theta\|^2 + \langle \theta, x^* \rangle > \langle \theta, x^* \rangle$$

$$= \max_{x \in K} \langle \theta, x \rangle$$

where we used that $y \notin K$ and hence $\|\theta\|^2 > 0$.

Therefore, we have shown that for any nonempty closed convex set K in $\mathbb{R}^n$ and any point y not in K, there exists a non-zero vector $\theta \in \mathbb{R}^n$ such that $\langle \theta, y \rangle > \max_{x \in K} \langle \theta, x \rangle$, which proves the Hyperplane Separation Theorem.

This theorem can be used to show that a polytope is essentially as general as a convex set. Any convex closed set can be written as the intersection of halfspaces. In other words, any convex set is a limit of a sequence of polyhedra.

Corollary

Any closed convex set K can be written as the intersection of halfspaces as follows

$$K = \bigcap_{\theta \in \mathbb{R}^n} \left\{ x : \ \langle \theta, x \rangle \leq \max_{y \in K} \langle \theta, y \rangle \right\}.$$

In other words, any closed convex set is a limit of a sequence of

polyhedra.

///// Proof /////

The Hyperplane Separation Theorem states that for any nonempty closed convex set K in $\mathbb{R}^n$ and any point y not in K, there exists a non-zero vector $\theta \in \mathbb{R}^n$ such that $\langle \theta, y \rangle > \max\limits_{x \in K} \langle \theta, x \rangle$.

This implies that the set of all points that do not belong to K, also called the complement of K, can be represented by the inequality $\langle \theta, x \rangle > \max_{y \in K} \langle \theta, y \rangle$.
Now, since the set of all vectors theta is a countable set, we can represent K as the intersection of all the halfspaces defined by the inequality above for all possible vectors θ.

Therefore, we have:

$$K = \bigcap_{\theta \in \mathbb{R}^n} \left\{ x : \ \langle \theta, x \rangle \leq \max_{y \in K} \langle \theta, y \rangle \right\}$$

Now, let L be defined as:

$$L = \bigcap_{\theta \in \mathbb{R}^n} \left\{ x : \ \langle \theta, x \rangle \leq \max_{y \in K} \langle \theta, y \rangle \right\}$$

Since all the points in K belong to all the halfspaces defined by the inequality above, it is clear that $K \subset L$.

Furthermore, for any point $x \notin K$, there exists a vector θ such that $\langle \theta, x \rangle > \max\limits_{y \in K} \langle \theta, y \rangle$, which means that x does not belong to the halfspace defined by this vector θ. Therefore, $x \notin L$, and so $L \subset K$.

Therefore, we have shown that $K = L$, and so any closed convex set can be represented as the intersection of all the halfspaces defined by the inequality above for all possible vectors θ, which completes the proof.