Convex Hulls

Convex sets are good, but good things are rare. When ugliness is inevitable, people usually wrap them by a layer of nicety. The same applies here, and our wrapping is named convex hull.

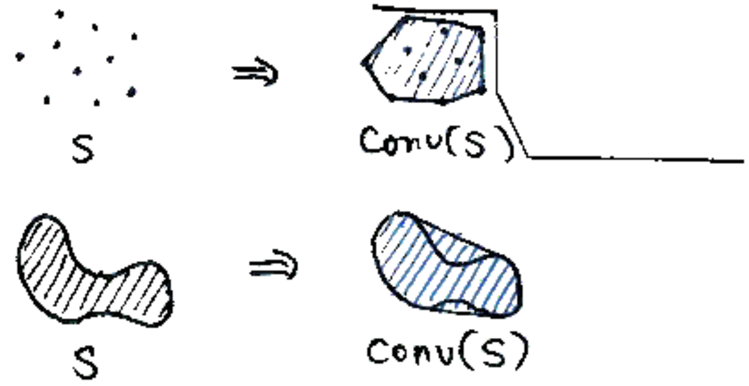
def convex hull.

The convex hull of a given set $S \subseteq \mathbb{R}^d$, denoted $\mathrm{Conv}(S)$, is the smallest (in the sense of set inclusion) convex set that contains S. In other words:

$$\mathrm{Conv}(S) := \bigcap_{\substack{C \text{ convex} \\ C \supseteq S}} C.$$

by Prop. 6, it's indeed convex, so $\mathrm{Conv}(S)$ exists.

e.g.



Pictorially, finding a convex hull of S is like shrinking a balloon that contains S. One may naturally wonder the other way round: what if we "expand" or "grow" a convex set from S?

def convex expansion.

The convex expansion of a set $S \subseteq \mathbb{R}^d$, denoted $\exp(S)$, is the set of all possible convex combinations of points in S. In other words:

$$\exp(S) := \left\{ \sum_{i=1}^{n} \lambda_i x_i \;\middle|\; n \in \mathbb{N},\ x_1, \cdots, x_n \in S,\ \sum_{i=1}^{n} \lambda_i = 1,\ \lambda_1, \cdots, \lambda_n \geq 0 \right\}.$$

Note that $\exp(S)$ is indeed convex: Take any two points, say $\sum_{i=1}^{m} \alpha_i x_i$ and $\sum_{i=1}^{m} \beta_i y_i$ from $\exp(S)$, it's easy to see that

$$\lambda \cdot \left(\sum_{i=1}^{n} \alpha_i x_i \right) + (1-\lambda)\left(\sum_{i=1}^{m} \beta_i y_i \right)$$

is again in $\exp(S)$, provided $\lambda \in [0,1]$. The argument is left as exercise.

We could regard $\exp(S)$ as the "most economic" way to make S convex: Recall def^n require that a convex set should be closed under convex combination, and that's exactly what $\exp(S)$ tries to fulfill.

So maybe not so surprisingly, the "shrink" and "growth" coincide:

Lemma 10.

$\mathrm{Conv}(S) = \exp(S)$.

Proof. ($\subseteq$): $\exp(S)$ is a convex set, which "participates in" the intersection that defined $\mathrm{conv}(S)$. So of course $\mathrm{conv}(S) \subseteq \exp(S)$.

($\supseteq$): $\forall \left(\sum_{i=1}^{n} \lambda_i x_i\right) \in \exp(S)$ must be in $\mathrm{Conv}(S)$ as well because

① $\mathrm{Conv}(S) \supseteq S \supseteq \{x_1, \cdots, x_n\}$

② $\mathrm{conv}(S)$ is convex; by def^n it must contain convex combination of $x_1, \cdots, x_n$. ■

Below we give yet another characterisation of convex hulls. It simplifies the convex sets that participate in the intersection.

Lemma 11. For $S \subseteq \mathbb{R}^d$, we have

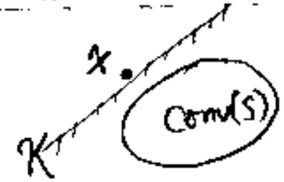
$$\mathrm{Conv}(S) = \bigcap_{\substack{H:\text{halfspace of } \mathbb{R}^d \\ H \supseteq S}} H.$$

In addition, if S is finite, then

$$\mathrm{Conv}(S) = \bigcap_{\substack{H:\text{ halfspace} \\ H \supseteq S \\ |H \cap S| \geq d}} H.$$

Proof. (1) General case: Obviously we have $\mathrm{Conv}(S) \subseteq \mathrm{RHS}$, so it remains to show $\mathrm{Conv}(S) \supseteq \mathrm{RHS}$.

Take a point $x \in \mathrm{RHS}$. Suppose to the contrary that $x \notin \mathrm{Conv}(S)$. Then the two convex sets $\{x\}$ and $\mathrm{Conv}(S)$ are disjoint, thus separated (not necessarily strictly) by a hyperplane K.



But then there is a halfspace (exactly the one side of K) that contains $\mathrm{conv}(S) \supseteq S$ but not containing x, contradicting that $x \in \mathrm{RHS}$.

(2) Finite case: the argument is very simple; observe that any halfspace $\mathcal{H} \supseteq S$: $|H \cap S| < d$ shall contain $(\mathcal{H}_1 \cap \mathcal{H}_2)$ for some $\mathcal{H}_1, \mathcal{H}_2 \supseteq S$: $|\mathcal{H}_1 \cap S| = |\mathcal{H}_2 \cap S| \geq d$.

■

Remark. Actually the proof of (1) is tricky: We use a stronger version of separation theorem than the one stated in Theorem 7. The technical issue is: $\mathrm{Conv}(S)$ might not be compact even if S is compact, so Theorem 7 is too weak. The stronger version is not too hard to prove, but has some technical details that we want to omit.

Lemma 11 tells us much about the shape of a convex hull when S is finite:

Corollary 12. For any finite $S \subseteq \mathbb{R}^2$, $\mathrm{Conv}(S)$ is a convex polygon whose vertices are from S.

Proof. Exercise. ■

In higher dimensions, we could also see that $\mathrm{Conv}(S)$ ($S \in \mathbb{R}^d$ finite) is bounded by finitely many hyperplanes. These shapes are "flat" in most parts and exhibits "edges" or "corners" somewhere. They are natural generalisations of convex polygons and are entitled "convex polytopes".

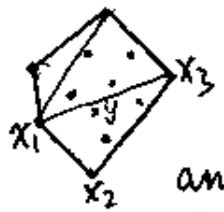
After understanding the basic shapes of convex hulls for finite sets, let's move on to present several more advanced properties of convex hulls. We start with a rather geometric and intuitive theorem by Carathéodory.

Theorem 13 (Carathéodory)
Suppose $S \subseteq \mathbb{R}^d$ with $n := |S|$. For any point $y \in \mathrm{Conv}(S)$, there exist $\leq d+1$ points from S, say $x_1, \cdots, x_{d+1}$, that

$$y \in \mathrm{conv}(\{x_1, \cdots, x_{d+1}\}).$$

In other words, only a few points are sufficient to "enclose" y.

Before proving it, we gain some intuition in the plane case $d=2$. As we already seen, $\mathrm{Conv}(S)$ is a convex polygon in this case.



So we could triangulate it and y lies in one of the triangles. Then we could choose x_1, x_2, x_3 to be the vertices of that triangle so that $y \in \mathrm{Conv}(\{x_1, x_2, x_3\})$.

So the basic idea of the proof is really to "triangulate" a convex polytope. Of course, there is no "triangles" in higher dimension. The analogue is simplex. In the induction step of the proof below, we "cut off" a simplex from the polytope ~~[illegible]~~ ~~[illegible]~~ and apply I.H. to the smaller polytope.

Proof.

Let us do induction on n. The base case $n \le d+1$ is vacuously true.

Now we do induction step $n \to n+1 \geq d+2$

Since $y \in \mathrm{Conv}(S) = \exp(S)$, we could write it as

$$y = \sum_{i=1}^{n+1} \lambda_i x_i \quad \cdots \circledast$$

Our goal is to move the weight of some λ_j onto other λ_i's (i.e. remove a simplex whose top is ~~[illegible]~~ x_j)

To this end, we consider the system of linear equations

$$\begin{cases} \sum_{i=1}^{n+1} \beta_i x_i = 0 & \Delta_1 \\ \sum_{i=1}^{n+1} \beta_i = 0 & \Delta_2 \end{cases} \cdots \Delta$$

where β_i's are unknowns. There are $(n+1)$ unknowns and $(d+1)$ equations (note that $\sum \beta_i x_i = 0$ is a vector equation). But $n+1 \geq d+2 > d+1$, so the system has a non-~~zero~~ trivial solution $\beta_1, \cdots, \beta_{n+1}$.

Now we ~~[illegible]~~ add $t\,\Delta_1$ to $\circledast$ and yield

$$y = \sum_{i=1}^{n+1} (t\beta_i + \lambda_i)\, x_i$$

for arbitrary $t \in \mathbb{R}$. The remaining task is easy: Choose $t := -\min_{i:\beta_i > 0} \frac{\lambda_i}{\beta_i}$, so that at least one coefficient $(t\beta_i + \lambda_i)$ becomes zero, while others remain non-negative, so the number of x_i's involved in expressing y is reduced by one. Finally, check that

$$\sum_{i=1}^{n+1} (t\beta_i + \lambda_i) = \sum_{i=1}^{n+1} \lambda_i = 1,$$

so I.H. applies. ■

Remark. Just to justify the intuition of Δ:

$\begin{cases} \text{(A}_1\text{): express some } \beta_j \text{ by combinations of other } \beta_i\text{'s} \\ \text{(A}_2\text{): ensure that the weight of } \beta_j \text{ is safely transferred to the weights of others, no more and no less.} \end{cases}$

Next we proceed to a theorem by Radon whose proof is short but involves clever algebraic tricks.

Theorem 14 (Radon)
Suppose $S \subseteq \mathbb{R}^d$ with $n := |S| \geq d+2$. Then we could partition S into $S^+ \cup S^-$ such that $\mathrm{Conv}(S^+) \cap \mathrm{Conv}(S^-) \neq \emptyset$.

Proof. Let $S = \{x_1, \cdots, x_n\}$. We extend each vector x_i by appending an extra coordinate:

$$\hat{x}_i := \begin{bmatrix} x_i \\ \hline 1 \end{bmatrix} \in \mathbb{R}^{d+1}.$$

Now consider the linear system

$$\sum_{i=1}^{n} \lambda_i \hat{x}_i = 0$$

where we have n unknowns $(\lambda_1, \cdots, \lambda_n)$ and $d+1 < d+2 \leq n$ equations. So there exists a nontrivial solution of $\lambda_1, \cdots, \lambda_n$. Then we could split them according to signs; formally,

$$I^+ := \{ i \in [n] : \lambda_i \geq 0 \} \neq \emptyset,$$
$$I^- := \{ i \in [n] : \lambda_i < 0 \} \neq \emptyset.$$

And we would derive

$$\begin{cases} \displaystyle\sum_{i \in I^+} \lambda_i x_i = \sum_{j \in I^-} (-\lambda_j) x_j & \cdots (1) \\ \displaystyle\sum_{i \in I^+} \lambda_i = \sum_{j \in I^-} (-\lambda_j) & \cdots (2) \end{cases}$$

due to the appended coordinate

This is almost done; note that (1) gives us a vector that could be expressed both by $\{x_i : i \in I^+\}$ and $\{x_j : j \in I^-\}$. The only issue is that it's not necessarily a <u>convex</u> combination. But if we use (2) to scale both sides, then the problem is resolved. Writing the idea down:

Divide (1) by (2), we have

$$\frac{\sum_{i\in I^+}\lambda_i x_i}{\sum_{i\in I^+}\lambda_i} = \frac{\sum_{j\in I^-}(-\lambda_j)x_j}{\sum_{j\in I^-}(-\lambda_j)} =: y$$

So clearly $y \in \exp(\{x_i : i \in I^+\})$ and at the same time $y \in \exp(\{x_j : j \in I^-\})$.

Hence from Lemma 10 we see

$$y \in \operatorname{conv}(\{x_i : i \in I^+\}) \cap \operatorname{Conv}(\{x_j : j \in I^-\}) \blacksquare$$

Remark. The proof immediately reminds us of the algebraic proof of a theorem about set systems: "If a set system $\mathcal{A} \subseteq 2^{[d]}$ has $n \geq d+2$ sets, then we could find disjoint subsystems $\mathcal{A}^-, \mathcal{A}^+ \subseteq \mathcal{A}$ such that

$$\bigcup_{A\in\mathcal{A}^-} A = \bigcup_{A\in\mathcal{A}^+} A \quad \text{and} \quad \bigcap_{A\in\mathcal{A}^-} A = \bigcap_{A\in\mathcal{A}^+} A."$$

Actually we could regard it as a corollary of Radon's theorem.

Radon's theorem has a generalised version, namely Tverberg's theorem:

Theorem 15 (Tverberg)

Suppose $S \subseteq \mathbb{R}^d$ with $n := |S| \geq (r-1)(d+1)+1$. Then we could partition S into disjoint $S_1, \dots, S_r$ such that $\bigcap_{i=1}^{r} \operatorname{Conv}(S_i) \neq \emptyset$.

The proof is still based on algebraic techniques, but we shall not expose it here.

Radon's theorem has some surprising consequences, for example the theorem below by Helly. It states that $(d+1)$-wise intersections of n convex sets in $\mathbb{R}^d$ is enough to enforce n-wise intersection.

Theorem 16 (Helly).

Given n convex sets $C_1, \dots, C_n \subseteq \mathbb{R}^d$ such that any $(d+1)$ of them intersect.

Then all of them intersect.

Proof. We proceed by induction on n. The base case $n = d+1$ is trivial.

Now we step from n to $n+1 \geq d+2$.
We will later use the key fact below:

> If we take multiple points from an intersection of convex sets, then the convex hull of these points still lie in the intersection. (why?)

Let us take $n+1$ points $x_1, \cdots, x_{n+1}$:

$$x_i \in \bigcap_{j \neq i} C_j \quad \text{(which is non-empty by I.H.)}$$

By Radon's theorem, we could partition $[n+1] = I^+ \cup I^-$ such that

$$\operatorname{conv}(X_{I^-}) \cap \operatorname{conv}(X_{I^+}) \neq \emptyset.$$

We take $y \in \operatorname{conv}(X_{I^-}) \cap \operatorname{conv}(X_{I^+})$ and observe what happens.

- By definition of the x_i's, we know
$$X_{I^-} \in \bigcap_{j \in I^+} C_j \quad \text{and} \quad X_{I^+} \in \bigcap_{j \in I^-} C_j.$$
- So of course
$$\operatorname{conv}(X_{I^-}) \subseteq \bigcap_{j \in I^+} C_j$$
$$\operatorname{conv}(X_{I^+}) \subseteq \bigcap_{j \in I^-} C_j$$
(by the fact on the left)
- Hence $y \in \bigcap_{j \in [n+1]} C_j$. ∎

To conclude this section, we give a nice application of Helly's theorem.
In one-dimension we have the notion of "medians", i.e. a point that could observe half of the point set from both sides. We naturally wonder if there is a similar notion in higher dimensions.

The generalised term is the "centerpoint".

def Centerpoint.
A centerpoint for a given point set $S \subseteq \mathbb{R}^d$ is any point $x \in \mathbb{R}^d$ such that
$\forall$ hyperplane passing through x,
both sides of the plane contains $\geq \frac{1}{d+1}|S|$ points from S.

e.g.

Note that we could not lift the constant $\frac{1}{d+1}$ to something like $\frac{1}{d}$; otherwise centerpoints clearly don't exist in many cases (for instance the top-right picture).

But even ~~so~~ with constant $\frac{1}{d+1}$, it is unclear whether centerpoints always exist. Proving existence would need judicious use of Helly's theorem.

Theorem 17 (Centerpoint theorem)
For any given finite pointset $S \subseteq \mathbb{R}^d$, there alway exists a centerpoint for S.

Proof. First we need a slightly different characterisation of centerpoints:

Claim x is centerpoint for S
$\iff$ $\forall$ halfspace $\mathcal{H}$: $|\mathcal{H} \cap S| > \frac{d}{d+1}|S|$ must contain x.

($\Rightarrow$) ~~If~~ Suppose there is some halfspace $\mathcal{H}$: $|\mathcal{H} \cap S| > \frac{d|S|}{d+1}$ that doesn't contain x, then the complement halfspace $\overline{\mathcal{H}}$ has $|\overline{\mathcal{H}} \cap S| < \frac{1}{d+1}|S|$ and $x \in \overline{\mathcal{H}}$, which is a contradiction to the definition of centerpoints.

($\Leftarrow$) Similar.

Now, we filter out ~~the~~ some subsets of S

$$\mathcal{A} := \{ \mathcal{H} \cap S : |\mathcal{H} \cap S| > \tfrac{d}{d+1}|S|,\ \mathcal{H} \text{ is a halfspace} \}$$

It suffices to show that

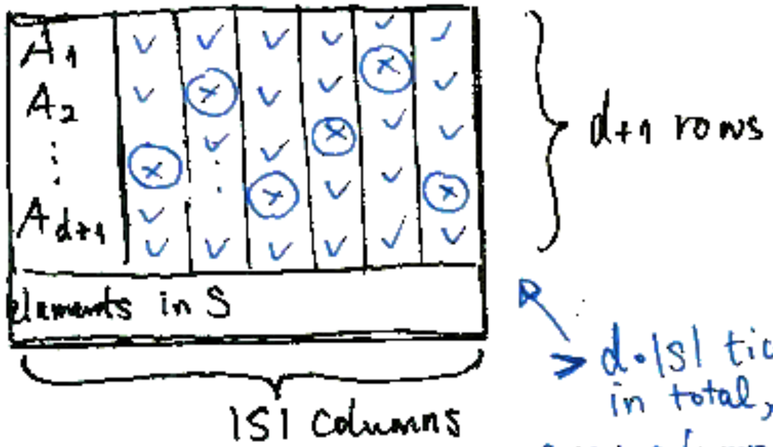
$$\bigcap_{A \in \mathcal{A}} \operatorname{conv}(A) \neq \emptyset .$$

To this end, we use Helly's theorem and reduce the proof to showing

$$\bigcap_{i=1}^{d+1} \operatorname{conv}(A_i) \neq \emptyset$$

for any $A_1, \dots, A_{d+1} \in \mathcal{A}$. But this is easy: By definition $|A_1|, \dots, |A_{d+1}| > \frac{d}{d+1}|S|$, so $\sum_{i=1}^{d+1} |A_i| > d \cdot |S|$.

Hence $A_1, \cdots, A_{d+1}$ must share at least one element.



$d \cdot |S|$ ticks needed in total, but if every column has to leave a vacancy then we could put $d \cdot |S|$ at most. ■