Convexity

- $\mathbb{R}^d$ deonotes the d-dimensional Euclidean space.
- A point in $\mathbb{R}^d$ is a d-tupple of real numbers, $x = (x_1, x_2, \ldots, x_d)$.
- Compare "linear" and "affine"

Linear Subspace

- a subset of $\mathbb{R}^d$ closed under addtion of vectors and under multiplication by real numbers
- Geometric meaning
  - In $\mathbb{R}^2$, the origin, all lines passing through the origin, and the whole of $\mathbb{R}^2$.
  - in $\mathbb{R}^3$, the origin, all lines passing through the origin, all planes passing through the origin, and the whole $\mathbb{R}^3$.

Linear Combination:
A linear combination of points, a_1, a_2, ..., $a_n \in \mathbb{R}^d$ is given by

$$\alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_n a_n,$$

where α_1, α_2, ..., α_n are real numbers.

Affine Subspace:
An affine subspace of $\mathbb{R}^d$ has the form $L + x$, where L is a linear subspace of $\mathbb{R}^d$, and x is a vector in $\mathbb{R}^d$.

- In $\mathbb{R}^3$, all points, all lines, all planes, and $\mathbb{R}^3$

Affine Combination

An affine combination of points a_1, a_2, ..., $a_n \in \mathbb{R}^d$ is given by

$$\alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_n a_n,$$

where α_1, α_2, ..., α_n are real numbers and $\alpha_1 + \alpha_2 + \ldots + \alpha_n = 1$

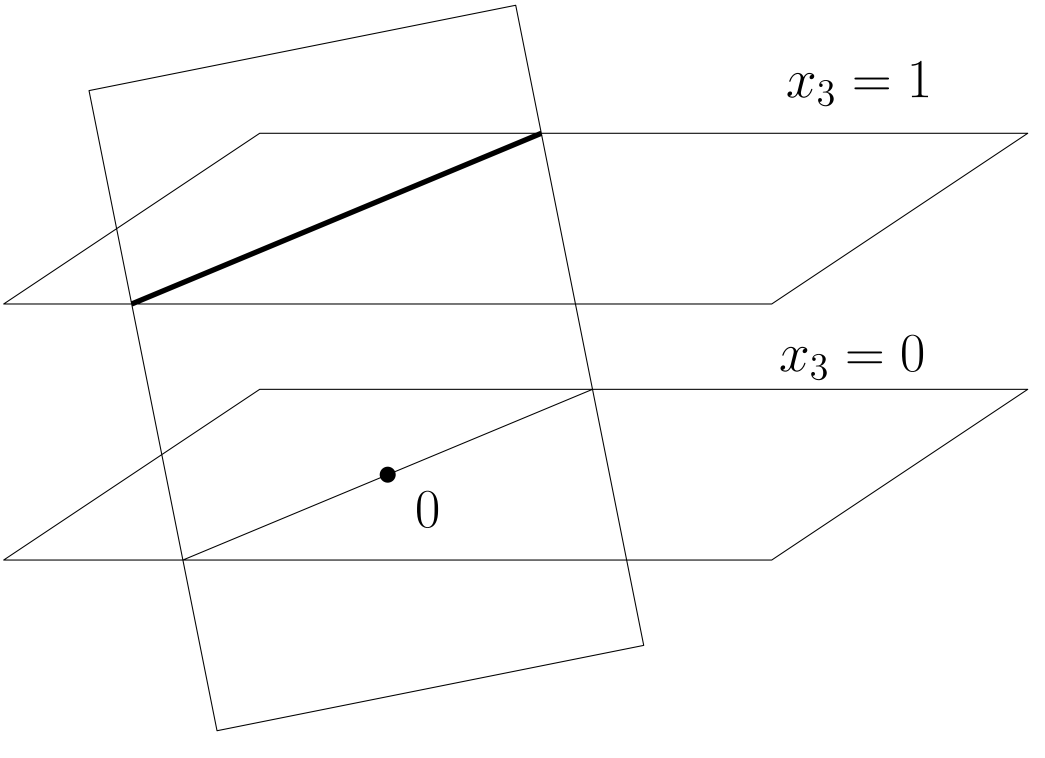
Linear Dependence

A set of points, $a_1, a_2, \ldots, a_n \in \mathbb{R}^d$, are ***linearly dependent*** if and only if there exists a set of real numbers, $\alpha_1, \alpha_2, \ldots, \alpha_n$, at least one of which is nonzero such that $\alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_n a_n = 0$

Affine Dependence

A set of points, $a_1, a_2, \ldots, a_n \in \mathbb{R}^d$, are ***affinely dependent*** if and only if there exists a set of real numbers, $\alpha_1, \alpha_2, \ldots, \alpha_n$, at least one of which is nonzero such that $\alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_n a_n = 0$ and $\alpha_1 + \alpha_2 + \ldots + \alpha_n = 0$.

Another viewpoint of affine dependence:
For a set of points, $a_1, a_2, \ldots, a_n \in \mathbb{R}^d$, let b_i be $(a_i, 1)$.
$a_1, a_2, \ldots, a_n$, are affinely dependent if and only if $b_1, b_2, \ldots, b_n$ are linearly dependent.



Additional viewpoint:
Let a_1, a_2, ..., a_{d+1} be points in $\mathbb{R}^d$, and let A be $d \times d$ matrix with $a_i - a_{d+1}$ i^{th} column for $1 \leq i \leq d$.
Then a_1, a_2, ..., a_{d+1} are affinely independent if and only if A has d linear independent columns, i.e., $\det(A) \neq 0$.
A useful criterion of affine independence using a determinant.

Names for affine subspaces:

- $(d-1)$-dimensional: hyperplane
- 0-dimensional: point
- 1-dimensional: line
- 2-dimensional: plane
- k-dimensional: k-flat

hyerplane

- usuall specified by a single linear equation of the form:

$$a_1x_1 + a_2x_2 + \cdots + a_dx_d = b.$$

  – the left hand side can be written as the scalar product $\langle a, x \rangle$

- expressed as

$$\{x \in \mathbb{R}^d \mid \langle a, x \rangle = b\},$$

  where $a \in \mathbb{R}^d \setminus \{0\}$ and $b \in \mathbb{R}$.

- A (closed) half-space in $\mathbb{R}^d$ is expressed as

$$\{x \in \mathbb{R}^d \mid \langle a, x \rangle \geq b\},$$

  – the hyerplane $\{x \in \mathbb{R}^d \mid \langle a, x \rangle = b\}$ is its boundary

k-flat:

- the intersection among $d - k$ hyperplanes
- viewed as a solution to a system

$$Ax = b$$

  of linear equestions, where $x \in \mathbb{R}^d$ is regarded as a column vector, A is a $(d-k) \times d$ matrix, and $b \in \mathbb{R}^{d-k}$.

General Position Assumption

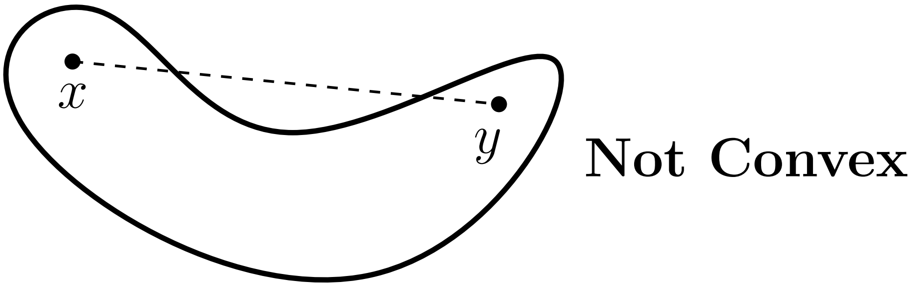
No $k+2$ points belongs to the same k-flat

- no three points are on the same line, no four points are one the plane, ...
- a magic phrase to avoid unlikely coincidence.

Convex Set

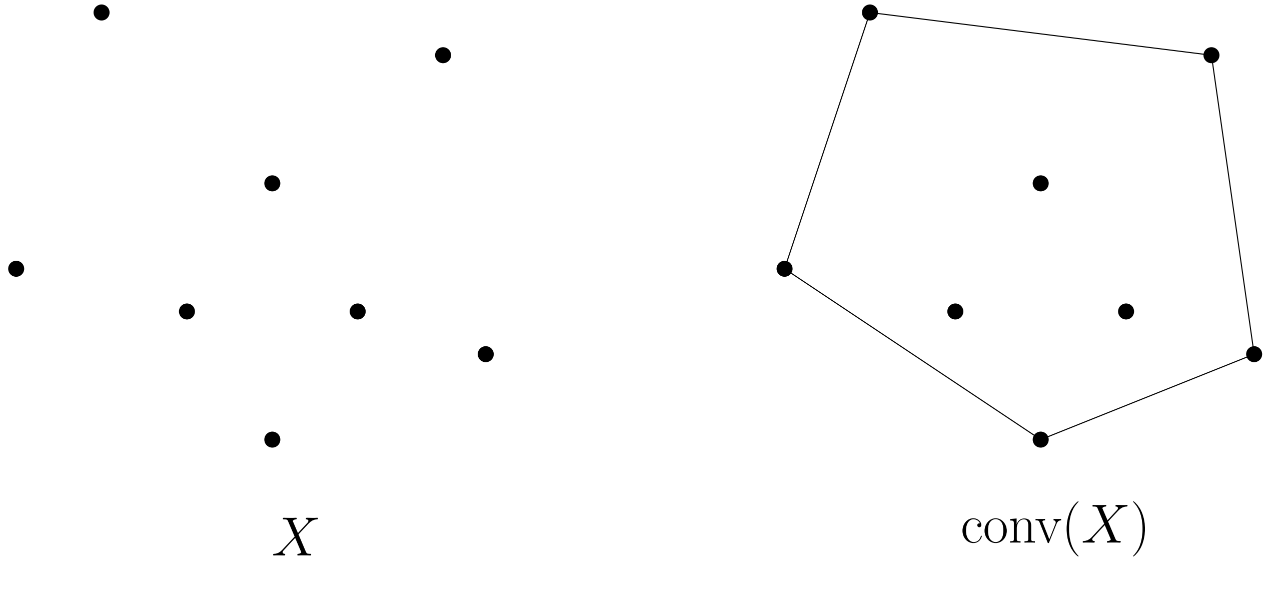
A set $C \subseteq \mathbb{R}^d$ is **convex** if for any two points $x, y \in C$, the whole segment $\overline{xy}$ is also contained in C. In other words, for every $t \in [0,1]$, the point $tx+(1-t)y$ belongs to C.

The intersection of an arbitrary familiy of convex sets is obviously convex.



Convex Hull

For a set $X \subseteq \mathbb{R}^d$, the **convex hull** of X, denoted by $\mathrm{conv}(X)$ is the intersection of all convex sets in $\mathbb{R}^d$ containing X.



Convex Combination

For points $x_1, \ldots, x_n \in \mathbb{R}^d$ and nonnegative real numbers $t_1, \ldots, t_n$, $t_1x_1 + t_2x_2 + \cdots + t_nx_n$ is a convex combination of $x_1, \ldots, x_n$ if $\sum_{i=1}^{n} t_i = 1$.

Lemma 1

A point x belongs to $\text{conv}(X)$ if and only if there exits points $x_1, x_2, \ldots, x_n \in X$ such that x is a convex combination of $x_1, x_2, \ldots, x_n$.

Sketch of proof

($\leftarrow$):

- Each convex combination of points of X must lie in $\text{conv}(X)$.
- For $n = 2$, it is true by definition.
- For large n, it can be proved by induction.

($\rightarrow$): Trivial

Important properties for convex huls regarding the dimensions d: Caratheodory's theorem, Radon's lemma, and Helly's theorem.

Caratheodory's Theorem

Let $X \subseteq \mathbb{R}^d$. Then each point of $\text{conv}(X)$ is a convex combination of at most $d + 1$ points of X.

Sketch of proof

- x is a convex combination of a finite number of points in X:

$$x = \sum_{i=1}^{k} t_i x_i,$$

  where $x_1, \ldots, x_k$ are points in X, $t_1, \ldots, t_k$ are nonnegative real numbers, and $\sum_{i=1}^{k} t_i = 1$
- It is sufficient to discuss $k > d + 1$.
- We will continuously remove one point until $k = d + 1$.

- Consider $k-1$ points/vectors $x_2 - x_1$, $x_3 - x_1$, $x_k - x_1$. Since they are linearly dependent, we have

$$\sum_{i=2}^{k} \lambda_i (x_i - x_1) = 0,$$

where λ_2, λ_3, ..., λ_k are real numbers.

- Let λ_1 be $-\sum_{i=2}^{k} \lambda_i$. Then we have

$$\sum_{i=1}^{k} \lambda_i x_i = 0 \text{ and } \sum_{i=1}^{k} \lambda_i = 0.$$

Note that at least one λ_i is larger than zero.

- For any real number α, we have

$$x = \sum_{i=1}^{k} t_i x_i + \alpha \sum_{i=1}^{k} \lambda_i x_i = \sum_{i=1}^{k} (t_i - \alpha \lambda_i) x_i.$$

- Let α be $\min_{1 \le i \le k} \{ \frac{t_i}{\lambda_i} \mid \lambda_i > 0 \}$ and let j be the index such that $\frac{t_j}{\lambda_j} = \alpha$
- Then $t_i - \alpha \lambda_i \ge 0$ for $1 \le i \le k$ and $t_j - \alpha \lambda_j = 0$.
- we can remove x_j

Application of Caratheodory's Theorem

- In $\mathbb{R}^2$, $\text{conv}(X)$ is the union of all triangles whose vertices are points in X.
- In $\mathbb{R}^3$, $\text{conv}(X)$ is the union of all tetrahedrons whose vertices are points in X.

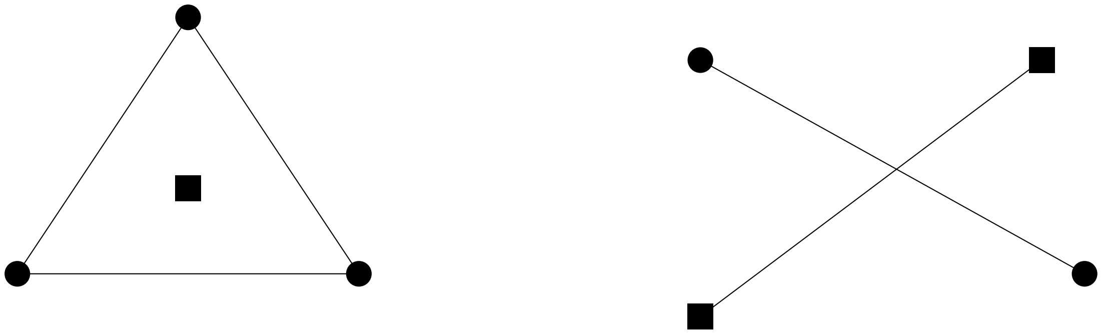
Radon's Lemma

Let A be a set of $d+2$ points in $\mathbb{R}^d$.
Then there exists two disjoints $A_1,\ A_2 \subset A$ such that

$$\operatorname{conv}(A_1) \cap \operatorname{conv}(A_2) \neq \emptyset.$$

Definition

A point $x \in \operatorname{conv}(A_1) \cap \operatorname{conv}(A_2)$, where A_1 and A_2 are as in Radon's Lemma, is called a **Radon point** of A, and the pair (A_1, A_2) is called a **Radon partition** of A. (it is easily seen taht we can require $A_1 \cup A_2 = A$.)



● A_1 ■ A_2

Sketch of Proof

- Let A be $\{a_1, a_2, \ldots, a_{d+2}\}$.
- These $d+2$ points are necessarily affinely dependent.
  $\rightarrow$ there exist real numbers $\alpha_1, \ldots, \alpha_{d+2}$, not all of which are 0, such that $\sum_{i=1}^{d+2} \alpha_i = 0$ and $\sum_{i=1}^{d+2} \alpha_i a_i = 0$.
- Let P be $\{i \mid \alpha_i > 0\}$ and N be $\{i \mid \alpha_i < 0\}$. It is clear that both P and N are nonempty.
- Let A_1 be $\{a_i \mid i \in P\}$ and A_2 be $\{a_i \mid i \in N\}$. We will exhibit a point x that is contained in both $\operatorname{conv}(A_1)$ and $\operatorname{conv}(A_2)$.
- Put S be $\sum_{i \in P} \alpha_i$; we also have $S = -\sum_{i \in N} \alpha_i$

- Then, define
$$x = \sum_{i \in P} \frac{\alpha_i}{S} a_i.$$
- Since $\sum_{i=1}^{d+2} \alpha_i a_i = 0 = \sum_{i \in P} \alpha_i a_i + \sum_{i \in N} \alpha_i a_i$, we have
$$x = \sum_{i \in N} \frac{-\alpha_i}{S} a_i.$$
- Since $\frac{\alpha_i}{S} \geq 0$ for $i \in P$ and $\sum_{i \in P} \frac{\alpha_i}{S} = 1$,
x is a convex combination of points in A_1, $\rightarrow x \in \mathrm{conv}(A_1)$.
- Since $\frac{-\alpha_i}{S} \geq 0$ for $i \in N$ and $\sum_{i \in N} \frac{-\alpha_i}{S} = 1$,
x is a convex combination of points in A_2, $\rightarrow x \in \mathrm{conv}(A_2)$.

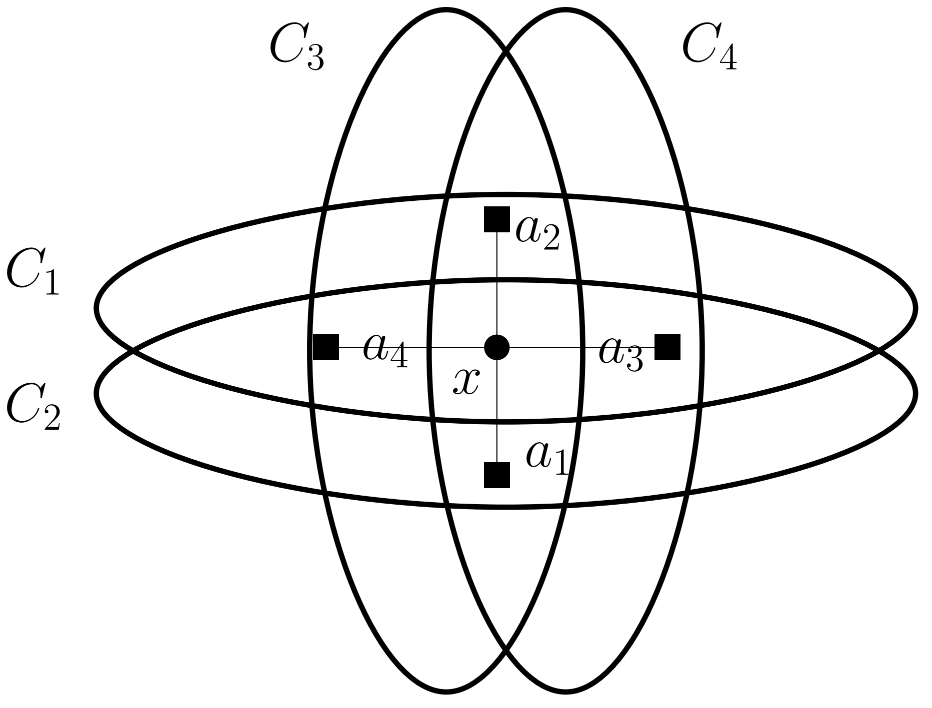
Helly's theorem
Let $C_1, C_2, \ldots, C_n$ be convex sets in $\mathbb{R}^d$, $n \geq d+1$.
If the intersection of every $d+1$ of these sets is nonempty, the intersection of all the C_i is nonempty.

Example
In $\mathbb{R}^2$, if every 3 among convex sets intersect, there is a point comon to all 100 sets.

Contrapositive viewpoint:
For $n \geq d+1$, if the intersection among $C_1, C_2, \ldots, C_n$ is empty, there exist $d+1$ sets among them whose intersection is empty.

Implication
For many planar problems, one can deal with 3 convex sets instead of an arbitrary number.

Wrong Form:
If every $d+1$ of convex sets intersect, all the n convex sets intersects. (Missing $n \geq d+1$).



Sktech of Proof (Using Radon's Lemma)

- For a fixed d, we prove by induction on n.
- The case $n = d + 1$ is trivial.
- Suppose that $n \geq d + 2$ and the statement holds for $n - 1$.
- Consider sets C_1, C_2, ..., C_n satisfying the assumption.
- If we leave out any one of these sets, the remaining sets have a nonempty intersection by the inductive assumption.
- Fix a point $a_i \in \bigcap_{j \neq i} C_j$ for $1 \leq i \leq n$ and consider a_1, a_2, ..., a_n.
- By Radon's Lemma, there exists index sets $I_1, I_2 \subset \{1, 2, \ldots, d+2\}$ such that
$$\text{conv}(\{a_i \mid i \in I_1\}) \cap \text{conv}(\{a_i \mid i \in I_2\}) \neq \emptyset.$$
- We pick a point x in the intersection, and prove that x lies in the intersection of all the C_i, leading to the statement.
- Consider some $i \in \{1, 2, \ldots, n\}$
- If $i \notin I_1$, each a_j with $j \in I_1$ lies in C_i, so $x \in \text{conv}(\{a_j \mid j \in I_1\}) \subseteq C_i$
- If $i \notin I_2$, each a_j with $j \in I_2$ lies in C_i, so $x \in \text{conv}(\{a_j \mid j \in I_2\}) \subseteq C_i$
- $x \in \bigcap_{i=1}^{n} C_i$.

An infinite version of Helly's theorem:
If we have infinite collection of convex sets in $\mathbb{R}^d$ such that any $d+1$ of them have a common point, the entire collection still need not have a coomon points. There are two examples in $\mathbb{R}^1$:

- $C_i = (0, 1/i)$ for $i = 1, 2, \ldots$
- $C_i = [i, \infty)$ for $i = 1, 2, \ldots$

The sets in the former are not closed, and those in the latter are unbounded.

Definition
A convex set is **compact** if it is closed and bounded

Infinite version of Helly's theorem
Let $\mathcal{C}$ be an arbitrary infinite famility of compact convex sets in $\mathbb{R}^d$ such that any $d+1$ of the sets have a nonempty intersection.
Then all the sets of $\mathcal{C}$ have a nonempty intersection.
Sketch of proof

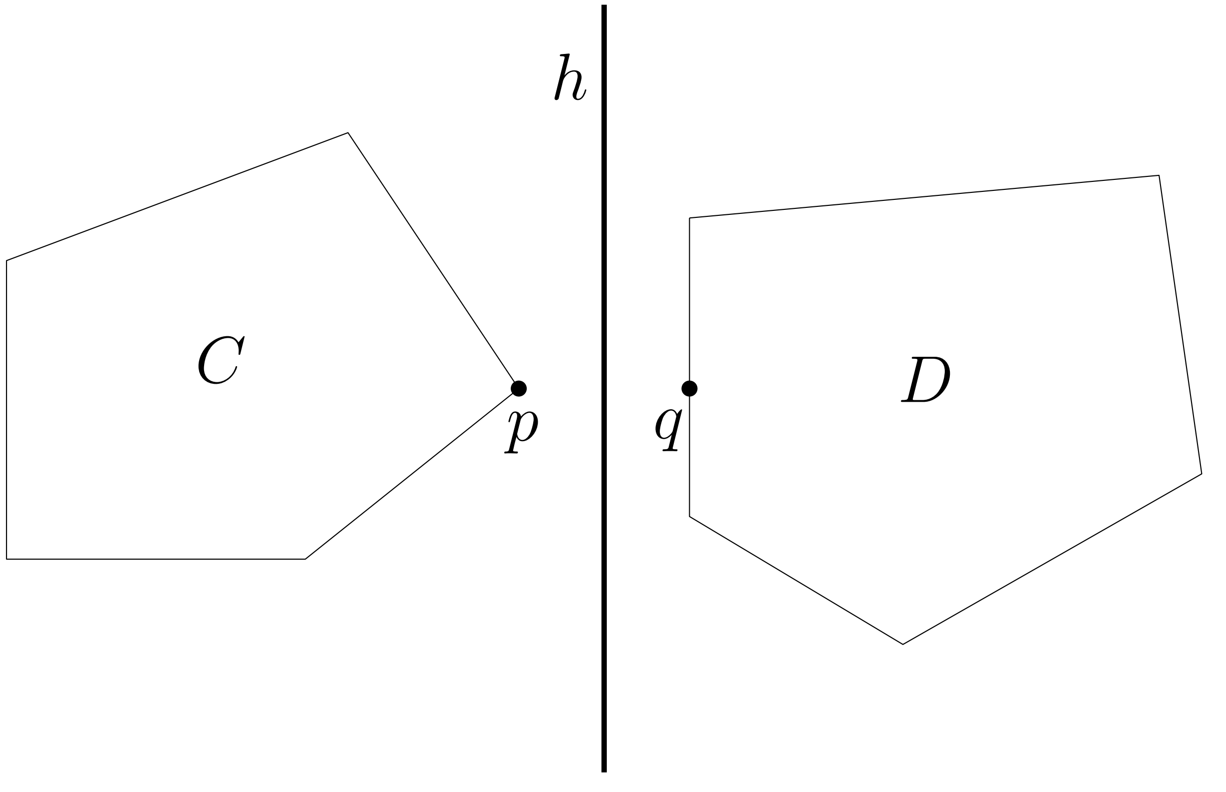
- By Helly's theorem, any finite subfamily of $\mathcal{C}$ has a nonempty intersection
- By a basic property of compactness, if we have an arbtrary family of compact sets such that each of its finite subfamilies has a nonempty intersection, the entire famility has a nonempty intesection.

Separation Theorem

Let $C, D \subseteq \mathbb{R}^d$ be convex sets with $C \cap D = \emptyset$.
If C and D are closed and at least one of them is bounded, there exists a hyperplane h such that C and D are separated by h, i.e., there exists a unit vector $a \in \mathbb{R}^d$ and a real number $b \in \mathbb{R}$ such that for all $x \in C$ $\langle a, x \rangle > b$, for all $x \in D$ $\langle a, x \rangle < b$, and h is $\{x \in \mathbb{R}^d \mid \langle a, x \rangle = b\}$.

Sketch of proof

- Since one of C and D is bounded, the distance between C and D is well-defined.
- Find $p \in C$ and $q \in D$ such that $d(p,q) = \min_{p' \in C, q' \in D} d(p', q')$, i.e., $d(p,q)$ is $d(C,D)$.
- h can be taken as the one perpdendicular to $\overline{pq}$ and passing through the midpoint of $\overline{pq}$.



Corollary

If both of C and D are unbounded and one of them is open, there exists a hyperplane h such that C lies in one of the closed half-spaces determined by h and D lies in the opposite closed half-space.

Farkas Lemma (one of many versions)
For $d \times n$ real matrix A, exactly one of the following cases occurs:

1. The system of linear equations $Ax = 0$ has a nontrivial nonnegative solution $x \in \mathbb{R}^n$. (All componennts of x are nonnegative and at least one of them is strictly positive.)
2. There exists a $y \in \mathbb{R}^d$ such that $y^T A$ is a vector with all entries strictly negative. Thus, if we multiply the j^{th} equation in the system $Ax = 0$ by y_j and add these equations together, we obtain an equation that obviously has no nontrivial nonnegative solution, since all the coefficients on the left-hand sides are strictly negative, while the right-hand side is 0.

Sketech of proof

- This is another version of separation theorem.
- Let $V \subset \mathbb{R}^d$ be the set of n points given by the column vectors of the matrix A.
- There two cases: either $0 \in \text{conv}(V)$ or $0 \notin \text{conv}(V)$.
- In the former case, we know that 0 is a convex combination of the points of V, and the coefficients of this convex combination determine a nontrivial nonnegative solution to $Ax = 0$.
- In the latter case, there exists a hyperplane strictly seprating V from 0, i.e., a unit vector $y \in \mathbb{R}^d$ such that $\langle y, v \rangle < \langle y, 0 \rangle = 0$ for each $v \in V$. This is just the y from the scond alternative in the Farkas lemma.

Definition (Center point)
Let X be an n-point set in $\mathbb{R}^d$.
A point $x \in \mathbb{R}^d$ is called a centerpoint of X if each closed half-space containing x contains at least $\frac{n}{d+1}$ points of X.

- X may generally have many centerpoints.
- A centerpoint need not belong to X.

The notion of centerpoint is a generalization of the median of one-dimensional data.

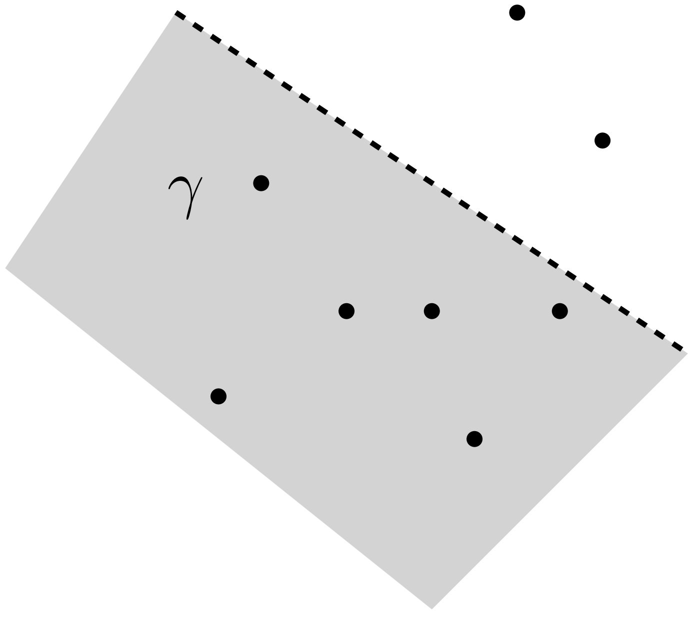
- suppose $x_1, \ldots, x_n \in \mathbb{R}$ are results of measurements of an unknown real parameter x.
- We can use the arithmetic mean, but if one of the measurement is complete wrong, the estimate is bad
- A median is more robust. It is a point x' such that at least $\frac{n}{2}$ of the x_i lies in the interval $(-\infty, x]$ and tat least $\frac{n}{2}$ of them lies in $[x, \infty)$.

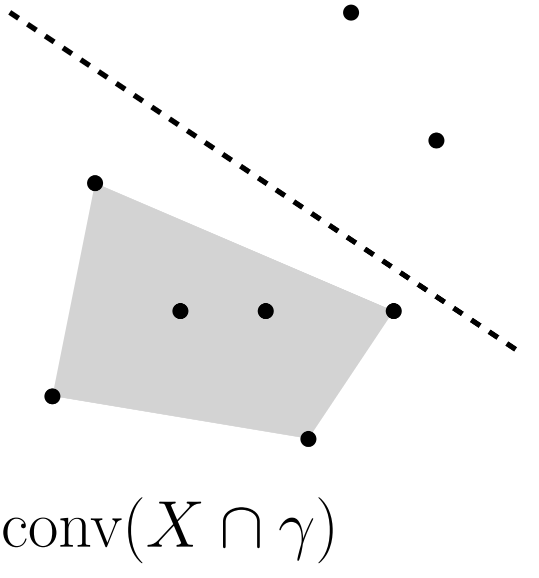
Definition (α-Center point)
Let X be an n-point set in $\mathbb{R}^d$.
A point $x \in \mathbb{R}^d$ is called an α-centerpoint of X if each closed half-space containing x contains at least α points of X. If $\alpha > \frac{1}{d+1}$, such a an α-centerpoint does not necessarily exist.

Cetnerpoint Theorem
Each finite point set in $\mathbb{R}^d$ has at least one centerpoint.

Sketch of proof

- An equivalent definition of a centerpoint:
  x is a centerpoint of X if and only if x lies in each open half-space γ such that $|X \cap \gamma| > \frac{d}{d+1}n$
- We attempt to apply Helly's theorem to conclude that all these open half-spaces intersect, i.e., x in the intersection.
- But they are open and unbounded, i.e., not compact.





- Instead of such an open half-space γ, we consider the compact convex set $\mathrm{conv}(X \cap \gamma) \subset \gamma$
- By letting γ run through all open half-spaces γ with $|X \cap \gamma| > \frac{d}{d+1}n$, we obtain a familiy $\mathcal{C}$ of compact convex sets.
- Each of those compact convex sets contains more than $\frac{d}{d+1}n$ points of X.
- The intersection of any $d+1$ of them contains at least one point of X.
- The family $\mathcal{C}$ consists of finitely many distinct sets since X has finitely many distinct subsets.
- By Helly's theorem, $\bigcap \mathcal{C} \neq \emptyset$.
- each point in $\bigcap \mathcal{C}$ is a centerpoint of X.