18.304 Talk II Convexity

1.1 linear and affine subspaces:

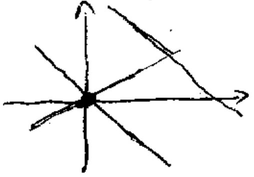
<u>linear subspace of $\mathbb{R}^d$</u>: a subset closed under addition of vectors & under multiplication by Real #s

Ex: $\mathbb{R}^3$: L.S. include: origin; all lines passing through ori
& all planes - - - - origin.

Affine notions: ex. in $\mathbb{R}^2$, General lines are Affine Subspaces

<u>Affine subspace of $\mathbb{R}^d$</u>: $x + L$

$\uparrow$ $x \in \mathbb{R}^d$ $\uparrow$ Linear subspace of $\mathbb{R}^d$



<u>Linear Combination</u> → linear span of a set X

<u>Affine Combination</u> of pts $a_1, \cdots a_n \in \mathbb{R}^d$

① Translate the whole set by $-a_n$ → a_n becomes the origin

② Make a linear combination & translate back by $+a_n$

$$\beta_1(a_1-a_n)+\beta_2(a_2-a_n)+\cdots+\beta_n(a_n-a_n)+a_n$$
$$=\beta_1 a_1+\beta_2 a_2+\cdots+\beta_{n-1}a_{n-1}+(1-\beta_1-\beta_2-\cdots-\beta_{n-1})a_n$$

where β_i are arbitrary numbers.

③ Affine Combination is of the form:

$$\alpha_1 a_1+\alpha_2 a_2+\cdots+\alpha_n a_n \quad \text{where } a_i \in \mathbb{R}$$
$$\& \ \alpha_1+\cdots+\alpha_n=1$$

Affine Hull : · · ·

Affine Dependence of pts $a_1, a_2 \cdots a_n$.

means one of them can be written as an affine combination of others :

exist real #s $\alpha_1, \alpha_2 \cdots \alpha_n$ at least one of them $\neq 0$

$$\left(\begin{array}{l} a_n = \alpha_1 a_1 + \cdots \ \alpha_{n-1} a_{n-1} \\ \alpha_1 a_1 + \cdots + \alpha_{n-1} a_{n-1} - a_n = 0 \\ \text{while } \alpha_1 + \cdots \ \alpha_{n-1} = 1 \text{, Therefore} \end{array} \right)$$

$$\alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_n a_n = 0 \ \& \ \alpha_1 + \alpha_2 + \cdots + \alpha_n = 0.$$

Affine dependence	Linear Dependence
$a_1, a_2 \cdots a_n$	$a_1 - a_n,\ a_2 - a_n \cdots a_{n-1} - a_n.$
Maximum Possible of Ind. pts in R^d	
$d+1$	d

Hyperplane : a $(d-1)$-dimensional affine subspace of $\mathbb{R}^d$ is called a hyperplane :

$$a_1 x_1 + a_2 x_2 + \cdots a_d x_d = b.$$

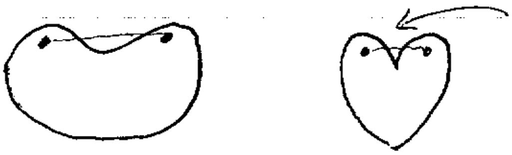
$$\{x \in \mathbb{R}^d,\ \langle a, x \rangle = b\}$$

a closed half-space in Rd is a set of the form $\Big)$ boundary

$$\{x \in \mathbb{R}^d : \langle a, x \rangle \geq b\} \text{ for some } a \in \mathbb{R}^d \setminus \{0\}$$

1.2. Convex sets, combination, separation.

Intuition:



Not allowed in a convex set.

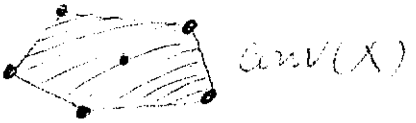
<u>Definition</u>: A set $C \subseteq \mathbb{R}^d$ is convex if for every two points $x, y \in C$ the whole segment xy is also contained in C:

for every $t \in [0,1]$, the pts $tx + (1-t)y \in C$.

<u>Convex Hull of $X \subseteq \mathbb{R}^d$</u>: $\mathrm{conv}(X)$ ← X doesn't have to be convex.

Intersection of all convex sets in $\mathbb{R}^d$ containing X.

ex:



<u>Convex Combination</u>:

for $x_1 \cdots x_n \in X$

$x = \cancel{\Sigma x}\sum_1^n t_i x_i$ s.t. $\sum_1^n t_i = 1$ ← nonnegative real #s.

<u>Claim</u>: $\mathrm{Conv}(X)$ = all <u>convex combination</u> of X.

Proof: ① Each convex combination of pts of X must lie in $\mathrm{conv}(X)$:

$n=2$ by definition.

$n=3$: $t_1 x_1 + t_2 x_2 + t_3 x_3$ ~~1-t~~ $1 - t_3 = t_1 + t_2$

$$= \underbrace{(1-t_3)\underbrace{\left(\frac{t_1}{t_1+t_2} x_1 + \frac{t_2}{t_1+t_2} x_2\right)}_{\in \mathrm{conv}(X)} + t_3 V_3}_{\in \mathrm{conv}(X)}$$

⋮

Same for n.

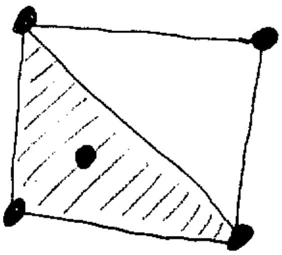
② Set of all convex combinations obviously contains X and it is convex.

Theorem (Carathéodory's theorem).

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

Proof:

Example: $\mathbb{R}^2$



Let x be a point in conv(X), then x:

$$x = \sum_{j=1}^{k} \lambda_j x_j \quad , \quad x_j \in X \, , \quad \lambda_j > 0 \ \& \ \sum \lambda_j = 1$$

Suppose ~~d~~ $k > d+1$ (otherwise, it's trivial). Then the points $x_2 - x_1, \cdots x_k - x_1$ are linearly dependent.

So there are real scalars $\mu_2, \cdots \mu_k$ not all zero st:

$$\sum_{j=2}^{k} \mu_j (x_j - x_1) = 0$$

x_i are affinely dependent $\left| \begin{array}{l} \sum \mu_j x_j = 0 \\ \sum \mu_j = 0 \end{array} \right.$

Let $\mu_1 = - \sum_{j=2}^{k} \mu_j \longrightarrow \left\{ \begin{array}{l} \sum_{1}^{k} \mu_j = 0 \\ \sum_{1}^{k} \mu_j x_j = 0. \end{array} \right.$

Not all $\mu_j = 0$, therefore, at least one $\mu_j > 0$. Then

$$x = \sum_{j=1}^{K} \lambda_j x_j - \alpha \underbrace{\sum_{j=1}^{K} \mu_j x_j}_{=0.} = \sum_{j=1}^{K} (\lambda_j - \alpha \mu_j) x_j$$

For any real α. In particular:

$$\alpha := \min_{1 \le j \le K} \left\{ \frac{\lambda_j}{\mu_j} : \mu_j > 0 \right\} = \frac{\lambda_i}{\mu_i} > 0$$

& $\lambda_j - \alpha \mu_j \geq 0$ $\left(\frac{\lambda_j}{\mu_j} - \alpha \geq 0. \right)$

In particular: $\lambda_i - \alpha \mu_i = 0$ by def. of α (i is the index of the smallest $\frac{\lambda_i}{\mu_i}$).

$$\therefore X = \sum_{j=1}^{K} \underbrace{(\lambda_j - \alpha \mu_j)}_{\text{non negative \& sum to 1}} x_j$$

(b/c. $\sum_{1}^{K} \alpha \mu_j = 0$.)

& $\lambda_i - \alpha \mu_i = 0$.

In other words, x is represented as a convex combination of at most $K-1$ points of P.

Repeat this process till x is represented as a convex combination of at most $d+1$ pts in P

Theorem (Hyperplane Separation Theorem)

Let $C, D \subseteq \mathbb{R}^d$ be convex sets with $C \cap D = \emptyset$.

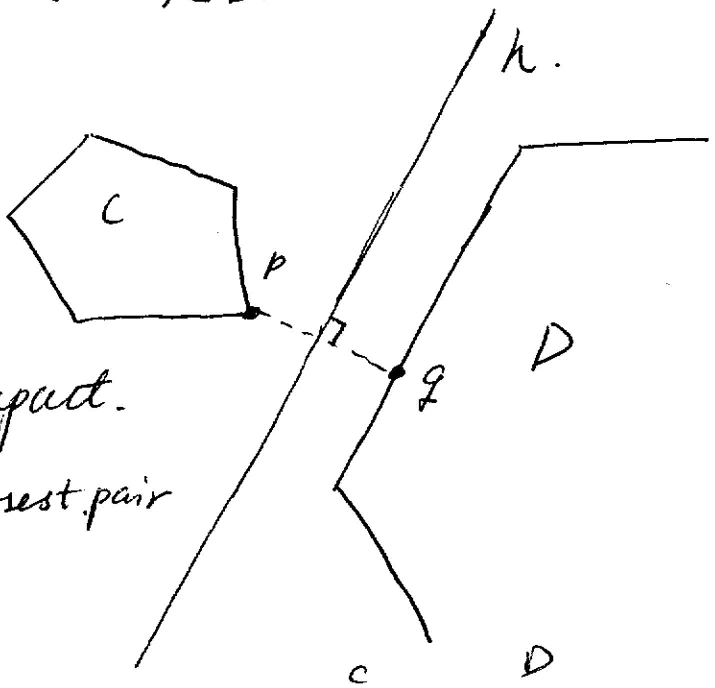
Then there exists a hyperplane h such that C lies in one of the closed half-spaces determined by h, & D lies in the opposite closed half-space.

In other words: there exists a unit vector $a \in \mathbb{R}^d$ & a number $b \in \mathbb{R}$ such that:

$$\forall x \in C \rightarrow \langle a, x \rangle \geqslant b$$

$$\& \ \forall x \in D \rightarrow \langle a, x \rangle \leq b.$$

Example: $\mathbb{R}^2$



i) Both are closed & at least one is compact.
Then you can find a closest pair (p, q).
Then any hyperplane perpendicular to the segment pq separate $\overset{C}{\cancel{A}}$ from $\overset{D}{\cancel{B}}$

ii) Spse C & D are arbitrary disjoint convex set
They can be exhausted by sequences of compact, convex subsets C_n & D_n $\rightarrow$ sequence of separating Hyperplane H_n.
$\hookrightarrow H$

Lemma (Farkas) :

For evey $d \times n$ real matrix A, exactly one of following

i). $Ax=0$ has a nontrivial nonnegative soln $x \in \mathbb{R}^n$

ii) $\exists\, y \in \mathbb{R}^d$ such that $y^T A$ is a vector with all entries strictly negative.

(Thus if we multiply jth Eqn in $Ax=0$ by y_j & add these eqns together, we obtain an eqn that has no nontrivial ~~soln~~ nonnegative soln)

Σ left = Negative Right=

Proof: (Another version of the separation Thm).

Let $V \subset \mathbb{R}^d$ be the set of n pts ~~det~~ given by the column vectors of the matrix A. There're two cases: ① $0 \in conv(V)$ ② $0 \notin conv(V)$

In case ①: 0 is a convex combination of pts of V & the coefficients of this convex ~~coefficient~~ combination determines a nontrivial nonnegative soln $Ax=0$

In Case ②: There exists a hyperplane strictly separating 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 $v \in V$.

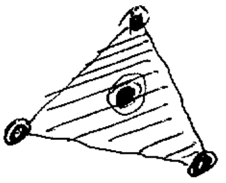
This is the y in the second case of Farkas lemma

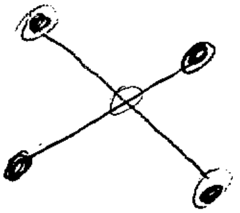
1.3 Radon's Lemma & Helly's Theorem:

Theorem (Radon). Let A be a set of d+2 pts in $\mathbb{R}^d$.
Then there exists two disjoint subsets $A_1, A_2 \subset A$ such that $conv(A_1) \cap conv(A_2) \neq \emptyset$.

(A point $x \in conv(A_1) \cap conv(A_2)$ is called *Radon Pt* of A
$\underbrace{(A_1, A_2)}_{\text{obv. require } A_1 \cup A_2 = A}$ is called a *Radon Partition* of A)

Proof: Ex:





Let $A = \{a_1, \dots a_{d+2}\}$. These d+2 pts are necessarily affinely dependent. $\exists\, \alpha_1 \cdots \alpha_{d+2} \in \mathbb{R}$ not all 0 such that

$$\sum_1^{d+2} \alpha_i = 0 \quad \sum_1^{d+2} \alpha_i a_i = 0$$

Set $P = \{i : \alpha_i > 0\}$, $N = \{i : \alpha_i < 0\}$. We claim P & N determine the desired subsets.

let $A_1 = \{a_i : i \in P\}$, $A_2 = \{a_i, i \in N\}$

Going to find a pt x that's in the convex Hull of both these sets.

$S = \sum_{i\in P} \alpha_i$, Also $S = -\sum_{i\in N} \alpha_i$

$$x = \sum_{i\in P} \frac{\alpha_i}{S} a_i$$

Since $\sum_{i=1}^{d+2} \alpha_i a_i = 0 = \sum_{i\in P} \underline{\alpha_i a_i} + \sum_{i\in N} \alpha_i a_i = \underset{\substack{\uparrow \\ x}}{\sum_{i\in P} \frac{\alpha_i a_i}{S}} + \sum_{i\in N} \frac{\alpha_i a_i}{S}$

$$x = \sum_{i\in N} \frac{-\alpha_i}{S} a_i$$

x is a convex combination of pts of A_1 & A_2.

<u>Theorem</u> (Helly's Theorem)

Let $C_1, C_2 \dots C_n$ be convex sets in $\mathbb{R}^d$, $n \geq d+1$.

Spse that the intersection of every $d+1$ of these sets is nonempty, then the intersection of all the C_i is nonempty. (Contraposition more useful.)

(Example)

<u>Proof</u> Induction on n: $n = d+1$ is clear

$n = d+2$ is the critical case, the rest follows by induction

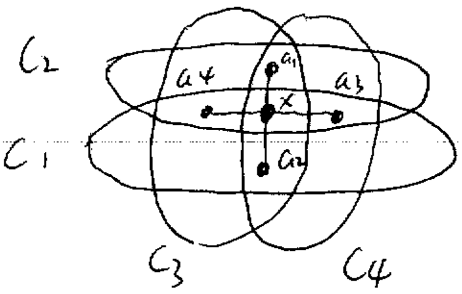
Consider sets $C_1, C_2 \dots C_n$ satisfying the assumption.

if we leave out one of these sets, the remaining sets have a nonempty intersection by inductive assumpt.

$n=4$
$d=2$

~~$\alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 = 1$~~

~~$\alpha_1 + \alpha_2 + \alpha_3 = 1$~~



Fix a point $a_i \in \bigcap_{j \neq i} C_j \Rightarrow a_1, a_2, \cdots a_{d+2}$

By Radon's lemma: there exists disjoint index set $I_1, I_2 \subset \{1, 2, \cdots d+2\}$ s.t.:

$$\mathrm{conv}(\{a_i : i \in I_1\}) \cap \mathrm{conv}(\{a_i : i \in I_2\}) \neq \emptyset$$

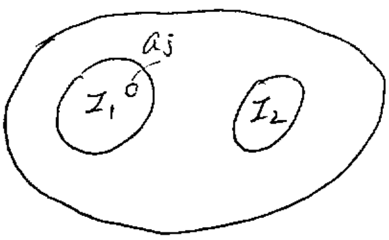
Pick pt x in this intersection

We claim x lies in the intersection of all C_i.

Consider some $i \in \{1, 2, \cdots n\}$. Then $\begin{cases} ① \; i \notin I_1 \\ ② \; i \notin I_2 \end{cases}$.

In the former case a_j with $j \in I_1$ lies in C_i

(if $j \in I_1$, then a_j is in all the set except C_j, $\rightarrow$ all c_i



$$\therefore \; x \in \mathrm{conv}(\{a_j : j \in I_1\}) \subseteq C_i$$

Similarly $x \in \mathrm{conv}(\{a_j : j \in I_2\}) \subseteq C_i$

$$\therefore \; x \in \bigcap_{i=1}^{n} C_i$$