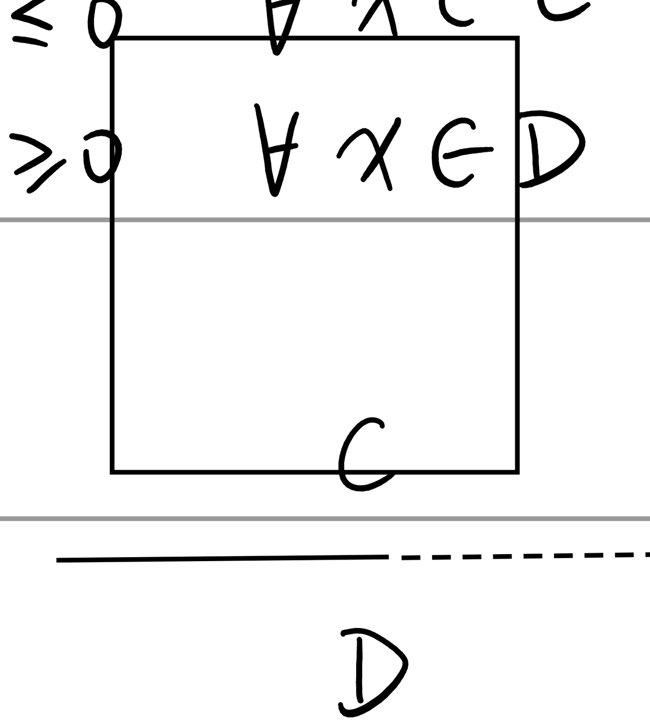
Lecture 5. Separating hyperplane theorem

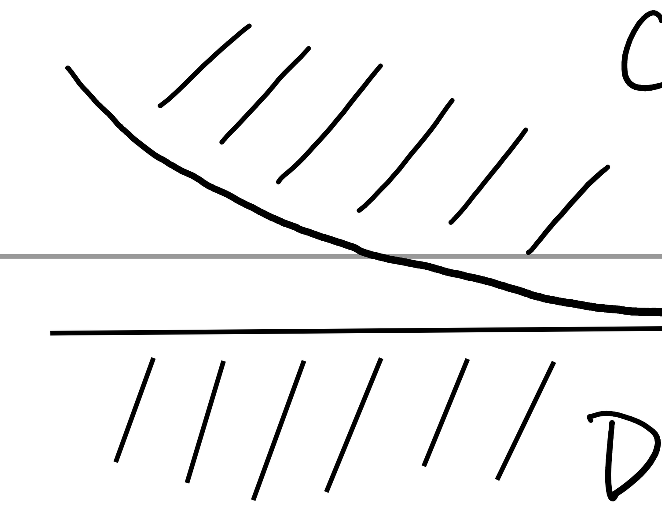
Separating hyperplane theorem: $\forall$ disjoint convex sets C, D. $\exists$ hyperplane $P: w^T x + b = 0$ separating C and D. i.e. $\begin{cases} w^T x + b \le 0 & \forall x \in C \\ w^T x + b \ge 0 & \forall x \in D \end{cases}$



Strict separating theorem.

Let C, D be two disjoint closed convex sets.

s.t. at least one of them is bounded.



Then $\exists$ strict separting hyperplane $w^T x + b = 0$.

Proof: Let $\text{dist}(C, D) \triangleq \inf_{x \in C, y \in D} \|x - y\|$. We claim that $\exists\, u \in C$ $v \in D$. s.t. $\|u - v\| = \text{dist}(C, D)$. Why? W.l.o.g. C is bounded given $u' \in C$, $v' \in D$. $\underbrace{D \cap \{x : \|u' - x\| \le \|u' - v'\|\}}_{\triangleq D'}$ $\underbrace{\text{bounded, closed}}_{\text{compact}}$.

$f_{u'}(x) = \|u' - x\|$ is continuous $\Rightarrow \exists\, \bar{v} = \arg\min_{x \in D'} f_{u'}(x)$. and $f_{u'}(\bar{v}) = \|u' - \bar{v}\| > 0$. Denote $\text{dist}(u', D) = \min_{x \in D'} f_{u'}(x)$ $\forall u' \in C$.

C is compact $\Rightarrow \exists\, \bar{u} = \arg\min_{x \in C} \text{dist}(x, D) \Rightarrow \exists\, \bar{u} \in C, \bar{v} \in D$.

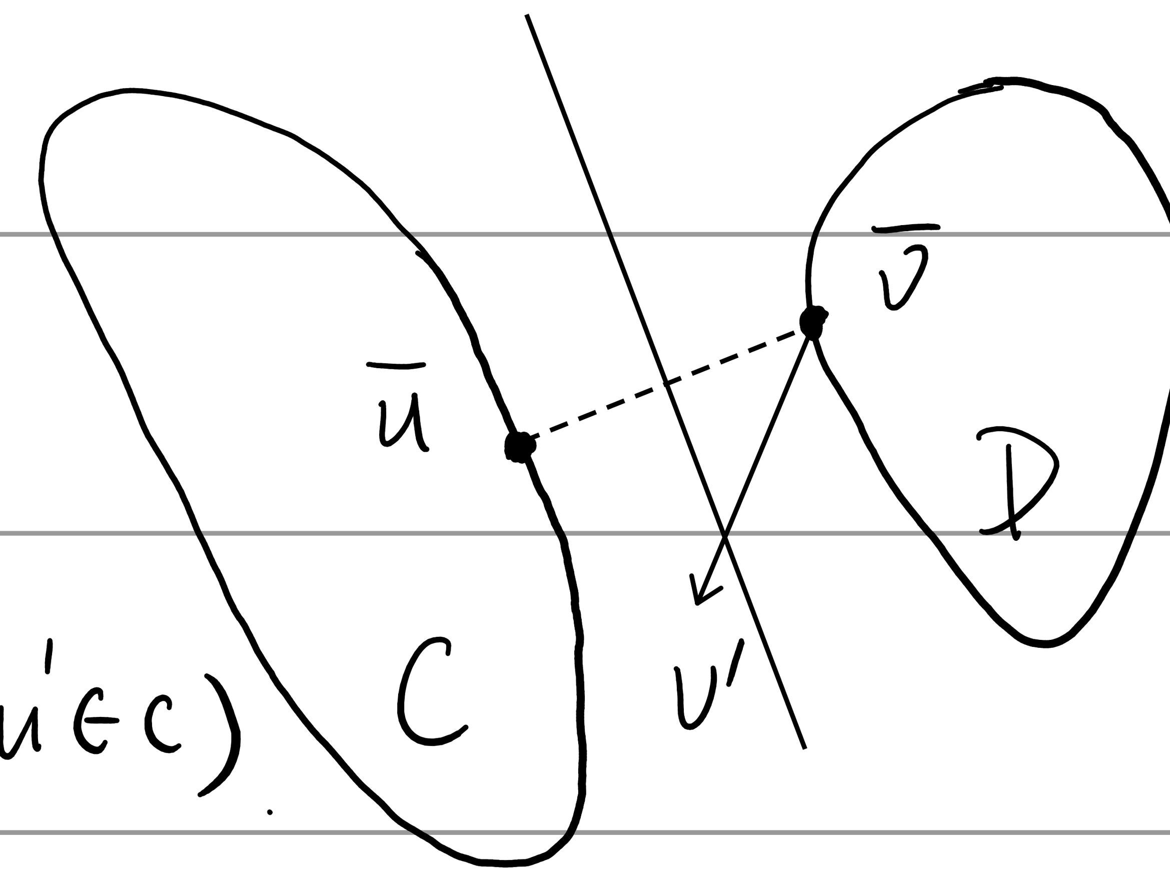
$\|\bar{u} - \bar{v}\| = \text{dist}(\bar{u}, D) = \text{dist}(C, D)$. Moreover, $\text{dist}(C, D) > 0$.

Let $w = \bar{v} - \bar{u}$, $b = \frac{1}{2}(\|\bar{u}\|^2 - \|\bar{v}\|^2)$. hyperplane $P: w^T x + b = 0$

We claim that P strictly separates C and D. Why set such b?

$w^T\left(\frac{\bar{u}+\bar{v}}{2}\right) + b = \frac{1}{2}(\bar{v}-\bar{u})^T(\bar{u}+\bar{v}) + \frac{1}{2}(\|\bar{u}\|^2 - \|\bar{v}\|^2) = 0$.

So P passes the mid point of $\bar{u}$ and $\bar{v}$.



Suppose for the sake of contradiction that

$\exists v' \in D$, s.t. $w^T v' + b \le 0$ (analogously if $\exists u' \in C$)

$\Rightarrow (\bar{v}-\bar{u})v' + \frac{1}{2}(\|\bar{u}\|^2 - \|\bar{v}\|^2) \le 0$. $v' \neq \bar{v}$ since $w^T\bar{v} + b =$

$(\bar{v}-\bar{u})^T\bar{v} + \frac{1}{2}(\|\bar{u}\|^2 - \|\bar{v}\|^2) = \frac{1}{2}\|\bar{u}-\bar{v}\|^2 > 0$. Define $f(x) = \|x-\bar{u}\|^2$.

Note that $f(x)$ achieves it minimum at $\bar{v}$ over the domain D.

By the first order necessary condition $\nabla f(\bar{v})^T d \ge 0$ $\forall$ feasible d.

$\bar{v} \in D$, $v' \in D$. by convexity of D, $d = v' - \bar{v}$ is a feasible direction.

$\nabla f(x) = \nabla (x-\bar{u})^T(x-\bar{u}) = 2(x-\bar{u})$. So $\nabla f(\bar{v})^T(v'-\bar{v}) =$

$2(\bar{v}-\bar{u})^T(v'-\bar{v}) = 2(\bar{v}^Tv' - \bar{u}v' - \|\bar{v}\|^2 + \bar{u}^T\bar{v})$ $\quad w^Tv' \le -b$.

$\le 2\left(\frac{1}{2}(\|\bar{v}\|^2 - \|\bar{u}\|^2) - \|\bar{v}\|^2 + \bar{u}^T\bar{v}\right) = (-\|\bar{v}\|^2 - \|\bar{u}\|^2 + 2\bar{u}^T\bar{v})$.

$= -\|\bar{u}-\bar{v}\|^2 < 0$. contradiction $\Rightarrow v' \neq D$. $\square$

Corollary 1. If $C \subseteq \mathbb{R}^n$. closed and convex. $d \in \mathbb{R}^n \setminus C$ is a point.

Then $\exists$ hyperplane P strictly separates C and d.

Corollary 2. supporting hyperplane theorem.

interior: $\text{int}\, C \triangleq \{x \in C : \exists \varepsilon > 0,\ \text{s.t. } \mathcal{B}(x, \varepsilon) \subseteq C\}$.

closure: $\text{cl}\, C \triangleq \{x \in \mathbb{R}^n : \exists x_1, \dots x_n, \dots \in C, \text{ s.t. } \lim_{n \to \infty} x_n = x\}$.

boundary: $\text{bd}\, C$, or $\partial C \triangleq \text{cl}\, C \setminus \text{int}\, C$, or equivalent defined as

$$\partial C \triangleq \{x \in \mathbb{R}^n : \forall \varepsilon > 0,\ \mathcal{B}(x, \varepsilon) \cap C \neq \phi, \text{ and } \mathcal{B}(x, \varepsilon) \nsubseteq C\}.$$

Remark: By definition $\text{int}\, C = C$ $\forall$ open C, and $\text{cl}\, C = C$. $\forall$ closed C.

Theorem (Supporting hyperplane theorem). $\forall C \neq \phi$ convex, $x_0 \in \partial C$

$\exists w \neq 0$, s.t. $\forall x \in C$, $w^T x \le w^T x_0$. hyperplane $P: w^T x - w^T x_0 = 0$

is called a supporting hyperplane of C at point x_0.

Proof. If $\text{int}\, C = \phi$. C lies in an affine set of dimension $< n$.

(o.w. $\exists$ $n+1$ affinely independent points in C, C contains an n-simplex).

An affine set lies on a hyperplane. It is a trivial supporting hyperplane.

If $\text{int}\, C \neq \phi$. let $C_{-\varepsilon} \triangleq \{x : \mathcal{B}(x, \varepsilon) \subseteq \text{cl}\, C\}$. By Corollary 1,

$\forall \varepsilon > 0$, $\exists w_\varepsilon \neq 0$ s.t. $w_\varepsilon^T x < w_\varepsilon^T x_0$, $\forall x \in C_{-\varepsilon}$. Normalize w_ε

s.t. $\|w_\varepsilon\| = 1$. Let $\varepsilon_k = \frac{1}{k} \to 0$. $\exists$ a subsequence of $\{w_{\varepsilon_k}\}$ has

a limit point w. We claim that w is a desired one. $\forall x \in \text{int}\, C$

$\exists N > 0$, s.t. $\forall k > N$. $w_{\varepsilon_k}^T x < w_{\varepsilon_k}^T x_0 \Rightarrow w^T x \le w^T x_0$.

$\forall y \in \partial C.\ \exists \{x_k \in C\} \to y.\ w^T x_k \le w^T x_0 \Rightarrow w^T y \le w^T x_0.$ $\square$.

Finally. we prove the separating hyperplane theorem: consider $C-D \triangleq \{u-v : u \in C, v \in D\}$. $C-D$ is also convex. It suffices to separates $C-D$ and $\{0\}$. i.e. find $w \ne 0$, s.t. $\forall x \in C-D$. $w^T x \le 0$.

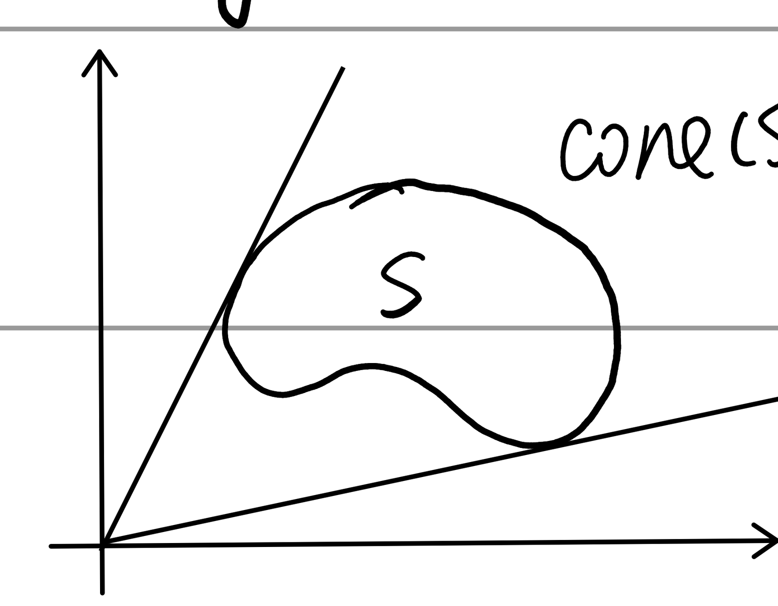
Case 1: $0 \notin \partial(C-D)$. strictly separating hyperplane theorem.

Case 2: $0 \in \partial(C-D)$. supporting hyperplane theorem at 0. $\square$.

Application: Farkas' lemma. $A \in \mathbb{R}^{n\times m}$. $b \in \mathbb{R}^n$. exactly 1 is true:

① $\exists x \in \mathbb{R}^m : Ax = b.\ x \geq \vec{0}$. ② $\exists y \in \mathbb{R}^n : A^T y \le 0.\ b^T y > 0$.

Proof. Recall the conic combination $\theta_1 x_1 + \cdots + \theta_k x_k$.



$\text{cone}(S) \triangleq \{\sum \theta_i x_i : \forall \theta_i \geq 0,\ x_i \in S\}$.

Let $A = (a_1, a_2, \dots, a_n)$. ① false $\Rightarrow b \notin \text{cone}(A)$. closed. convex

By separating hyperplane (strict). $\exists y, z$. $\begin{array}{ll} u^T y + z < 0 & \forall u \in \text{cone}(A) \\ b^T y + z > 0 & \end{array}$

$\forall a_i.\ \forall \lambda_i \geq 0.\ \lambda_i a_i^T y + z < 0 \Rightarrow a_i^T y + \frac{z}{\lambda_i} < 0 \Rightarrow a_i^T y \le 0$.

$0 \in \text{cone}(A) \Rightarrow z < 0 \Rightarrow b^T y > 0$. so y is a desired one. $\square$

More about separating and supporting hyperplanes.

① Why did we emphasis "finite" halfspaces when defining polytope?

Proposition: $\forall$ closed convex sets is the intersection of halfspaces.

Proof. Consider all hyperplanes $\{P: w^T x = w^T x_0 : x_0 \in \partial C\}$.

Let D be the intersection of all halfspaces: $w^T x \leq w^T x_0$ or $\geq$

$C \subseteq D$ is trivial (by definition). We now prove $D \subseteq C$. If $\exists u \notin C$.

$u \in D$. let $v = \arg\min_{x \in C} \|x - u\|$ so. $\text{dist}(u, C) = \|u - v\|$.

Consider the proof of strict separating hyperplane w.r.t u and C,

$w^T v' \leq -b$. $b = \bar{u}^T \bar{v} - \|\bar{v}\|^2$ suffices. $w^T \bar{v} + b = 0$. So it passes

v, thus is a supporting hyperplane of C, but separates u and C

$\Rightarrow w^T v' + b \geq 0$, $\forall v' \in C$. and $w^T u + b < 0 \Rightarrow u \notin D$. $\square$

② converse theorem of separating hyperplanes: $\left.\begin{matrix}\text{convex}\\ \text{disjoint}\end{matrix}\right\} \Rightarrow$ separatable

Theorem. $\forall$ convex C. D. ≥ 1 open. disjoint iff $\exists$ separating hyperplanes.

Proof: "$\Leftarrow$". w.l.o.g assume C is open. and $P: w^T x + b = 0$. separating.

$w^T x + b \begin{matrix} \leq 0 & \forall x \in C \\ \geq 0 & \forall x \in D. \end{matrix}$ if $C \cap D \neq \phi$. $\exists v \in C \cap D$. $w^T v + b = 0$.

C is open $\Rightarrow \exists \varepsilon > 0$. $B(x, \varepsilon) \subseteq C \Rightarrow \exists \overset{v + \frac{\varepsilon}{2} w}{\overset{\parallel}{u}} \in C$. $w^T u + b > 0$. $\square$

③ converse theorem of supporting hyperplanes: convex $\Rightarrow$ supportable.

Theorem. $\forall$ closed C. int $C \neq \phi$. $\forall x_0 \in \partial C$ supportable $\Rightarrow C$ convex.