IOE 611/ Math 663
Lecture 3: Convex sets (2)

Salar Fattahi

September 6, 2022

Agenda

- Operations that preserve convexity (cont'd)
- Separating and supporting hyperplane theorems

Recap (1)

- **cone**: a set C such that for any $x \in C$ and $\theta \geq 0$, $\theta x \in C$
  convex cone: a cone that is convex, *i.e.*, set that contains all conic combinations

$$\{\theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_k x_k \mid \theta_i \geq 0\}$$

  of points in the set (i.e., closed under conic combinations)
- **Positive semidefinite cone**: $\mathbf{S}_+^n = \{X \in \mathbf{S}^n \mid X \succeq 0\}$: positive semidefinite $n \times n$ matrices

$$X \in \mathbf{S}_+^n \quad \Longleftrightarrow \quad z^T X z \geq 0 \text{ for all } z$$

Recap (1)

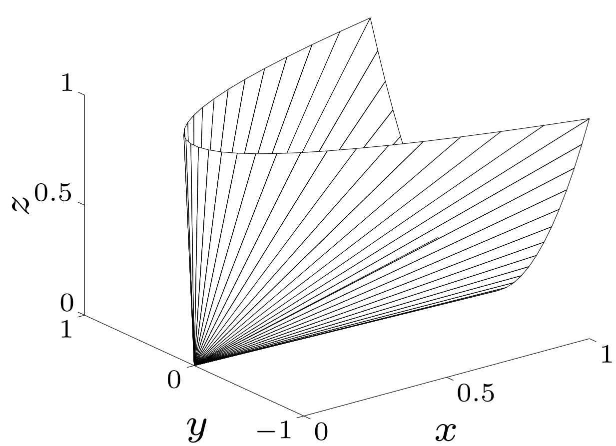
- **cone**: a set C such that for any $x \in C$ and $\theta \geq 0$, $\theta x \in C$
  convex cone: a cone that is convex, *i.e.*, set that contains all conic combinations

$$\{\theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_k x_k \mid \theta_i \geq 0\}$$

  of points in the set (i.e., closed under conic combinations)
- **Positive semidefinite cone**: $\mathbf{S}_+^n = \{X \in \mathbf{S}^n \mid X \succeq 0\}$: positive semidefinite $n \times n$ matrices

$$X \in \mathbf{S}_+^n \quad \Longleftrightarrow \quad z^T X z \geq 0 \text{ for all } z$$

example: $\begin{bmatrix} x & y \\ y & z \end{bmatrix} \in \mathbf{S}_+^2$



Recap (2)

1. Cartesian product
   ($S_1 \times S_2 = \{(x_1, x_2) \mid x_1 \in S_1,\ x_2 \in S_2\}$; if S_1 and S_2 are convex, so is $S_1 \times S_2$)
2. intersection
3. the image of a convex set under $f(x) = Ax + b$ is convex

$$S \subseteq \mathbf{R}^n \text{ convex} \quad \Longrightarrow \quad f(S) = \{f(x) \mid x \in S\} \subseteq \mathbf{R}^m \text{ convex}$$

Affine function

suppose function $f : \mathbf{R}^n \to \mathbf{R}^m$ is affine ($f(x) = Ax + b$ with $A \in \mathbf{R}^{m \times n}$, $b \in \mathbf{R}^m$)

- the inverse image $f^{-1}(C)$:

$$f^{-1}(C) = \{x \in \mathbf{R}^n \mid f(x) \in C\}$$

the inverse image of a convex set under f is convex

$$C \subseteq \mathbf{R}^m \text{ convex} \quad \Longrightarrow \quad f^{-1}(C) \text{ convex}$$

All the Points $x \in \mathbb{R}^n$ that lead's to

$f(x) \in C$

Examples

LMI

- **example:** solution set of linear matrix inequality
  $\{x \mid x_1 A_1 + \cdots + x_m A_m \preceq B\}$ (with $A_i, B \in \mathbf{S}^p$)

LMI appear in lot of field's, control theory, Robotic's,

$A, B \in S^n$ (Symmetric, n dimention)

$A \preceq B$: $B - A \in S_+^n \succeq 0$ (PSD)

The set of all Possible x's such that

$x_1 A_1 + \cdot\cdot + x_m A_m \preceq B$

$A_i \; \forall i, \; B \in S^n$

Such that $B - A_i \in S^n_+$ (PSD)

$\Rightarrow B - A_i \succeq 0$

let $f(x_i) = B - x_1 A - x_2 A_2 - x_3 A_3 - \cdots - x_m A_m$

here x is a vector $\Rightarrow x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \in \mathbb{R}^m$

$\Rightarrow f(x) = B - x_1 A_1 - x_2 A_2 - \cdots - x_m A_m$

Linear function of element's of x

$\Rightarrow f(x)$ is a Affine function.

Let's define $C = S^n_+$ (set of all PSD cone's)

Then $f^{-1}(S_n^+) = \{x : f(x) \in S^n_+\}$

$= \{x : B - x_1 A_1 - x_2 A_2 \cdots - x_m A_m \in S^n_+\}$

$= \{x : x_1 A_1 + x_2 A_2 \cdots + x_m A_m \preceq B\}$

This set is inverse image of S^n_+ (PSD) which is convex $\Rightarrow$ The set x must be convex

$\Rightarrow \{x : x_1 A_1 + x_2 A_2 + \cdots + x_m A_m \preceq B\}$

is convex.

Perspective and linear-fractional function

perspective function $P : \mathbf{R}^{n+1} \to \mathbf{R}^n$:

$$P(x,t) = x/t, \qquad \mathbf{dom}\, P = \{(x,t) \mid t > 0\}$$

images and inverse images of convex sets under perspective are convex

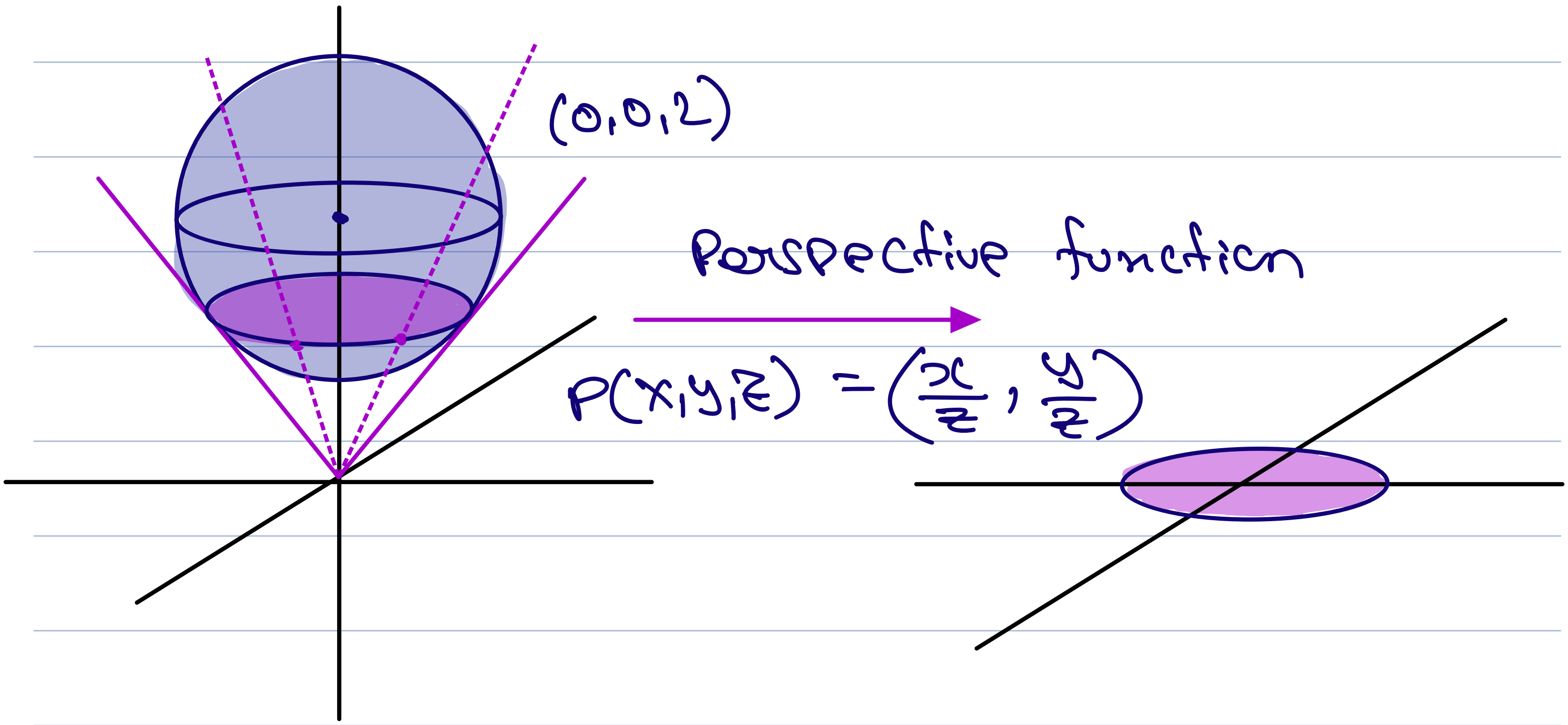
$$P(x,t) = \left[\frac{x_1}{t}, \frac{x_2}{t}, \frac{x_3}{t}, \dots, \frac{x_n}{t}\right]$$

if $S \in \mathbf{dom}\, P$ is convex, then $\left\{\frac{x}{t} : (x,t) \in S\right\}$ is also convex.

* The inverse image of C under perspective function

$\Rightarrow$ if C is convex, then $\left\{(x,t) : \frac{x}{t} \in C, t > 0\right\}$ is also convex

$x^2+y^2+(z-2)^2=1$



why do we care about Perspective function's?

* it won't look useful, But if we combine it with other Properties (operation's that preserve convexity) we can prove the convexity of set's that are not very trivial

(Ex: Linear – fractional function)

Perspective and linear-fractional function

perspective function $P : \mathbf{R}^{n+1} \to \mathbf{R}^n$:

$$P(x, t) = x/t, \qquad \mathbf{dom}\, P = \{(x, t) \mid t > 0\}$$

images and inverse images of convex sets under perspective are convex

linear-fractional function $f : \mathbf{R}^n \to \mathbf{R}^m$:

$$f(x) = \frac{Ax + b}{c^T x + d}, \qquad \mathbf{dom}\, f = \{x \mid c^T x + d > 0\}$$

images and inverse images of convex sets under linear-fractional functions are convex

$x \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$, $d \in \mathbb{R}$

$\Rightarrow \quad Ax+b \in \mathbb{R}^m, \quad c^T x + d \in \mathbb{R}$

$\Rightarrow \quad \dfrac{Ax+b}{c^T x + d} \in \mathbb{R}^m$

$\Rightarrow \quad f(x) = \dfrac{Ax+b}{c^T x + d} \Rightarrow f(x) : \mathbb{R}^n \longrightarrow \mathbb{R}^m$

$\mathrm{dom}\, f = \{ x \in \mathbb{R}^n \mid c^T x + d > 0 \}$ is convex

$x \in \mathbb{R}^n$ (convex) $\Rightarrow$ $Ax+b$ is convex (Affine function)

$c^T x + b$ is convex (Affine function)

$\Rightarrow C_1 = \{ (Ax+b, c^T x + b) : x \in \mathbb{R}^n \}$ convex

(Cartesian Product)

$\Rightarrow \quad C_2 = \{ (Ax+b, c^T x + b) : x \in \mathbb{R}^n, c^T x + b > 0 \}$ convex

intersection of two convex set's

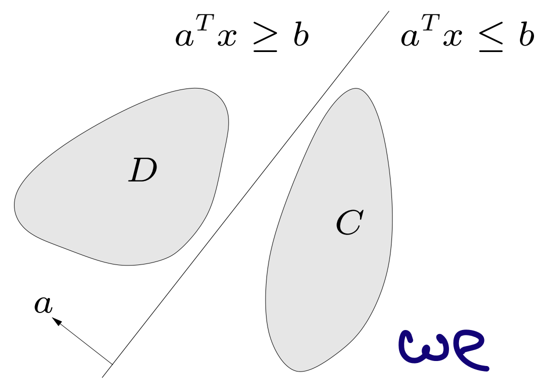
$\Rightarrow \quad C_3 = \left\{ \dfrac{Ax+b}{c^T x + b} ; \; x \in \mathbb{R}^n, \; c^T x + d \geq 0 \right\}$

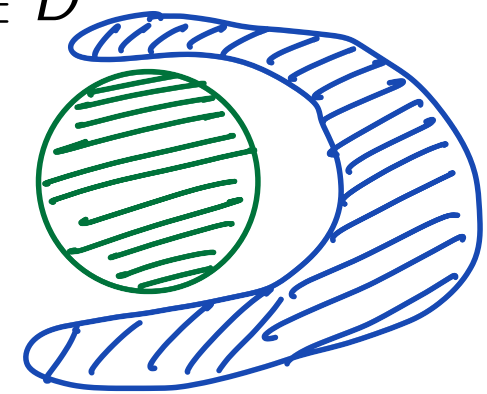
Convex: (Image of Perspective function

Separating hyperplane theorem

if C and D are disjoint convex sets, then there exists $a \neq 0$, b such that

$$a^T x \leq b \text{ for } x \in C, \qquad a^T x \geq b \text{ for } x \in D$$





we cannot find hyperplane

the hyperplane $\{x \mid a^T x = b\}$ separates C and D

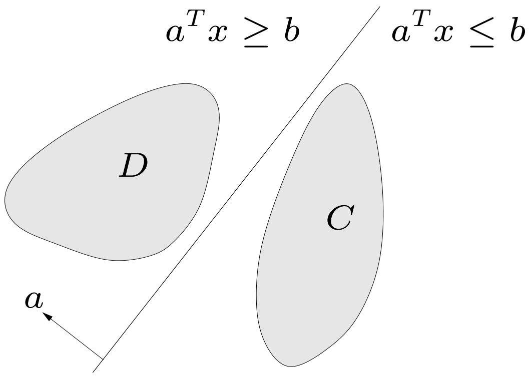
$\Rightarrow$ Probably one of the most fundamental theorem in optimization.

$\Rightarrow$ A Lot of theorem's are gonna be BAsed on this theorem.

Separating hyperplane theorem

if C and D are disjoint convex sets, then there exists $a \neq 0$, b such that

$$a^T x \leq b \text{ for } x \in C, \qquad a^T x \geq b \text{ for } x \in D$$



the hyperplane $\{x \mid a^T x = b\}$ separates C and D

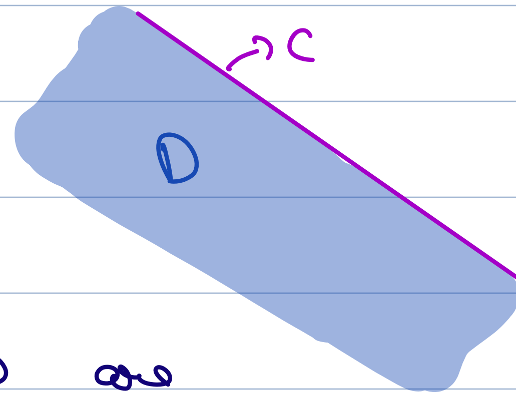
strict separation ($a^T x < b$ for $x \in C$, $\quad a^T x > b$ for $x \in D$) is not always possible.

Ex: Give an example of two convex sets that are disjoint, yet we cannot find a hyperplane that seperates strictly.

(1)

$C = \{x : a^T x = b\}$

$D = \{x : a^T x < b\}$



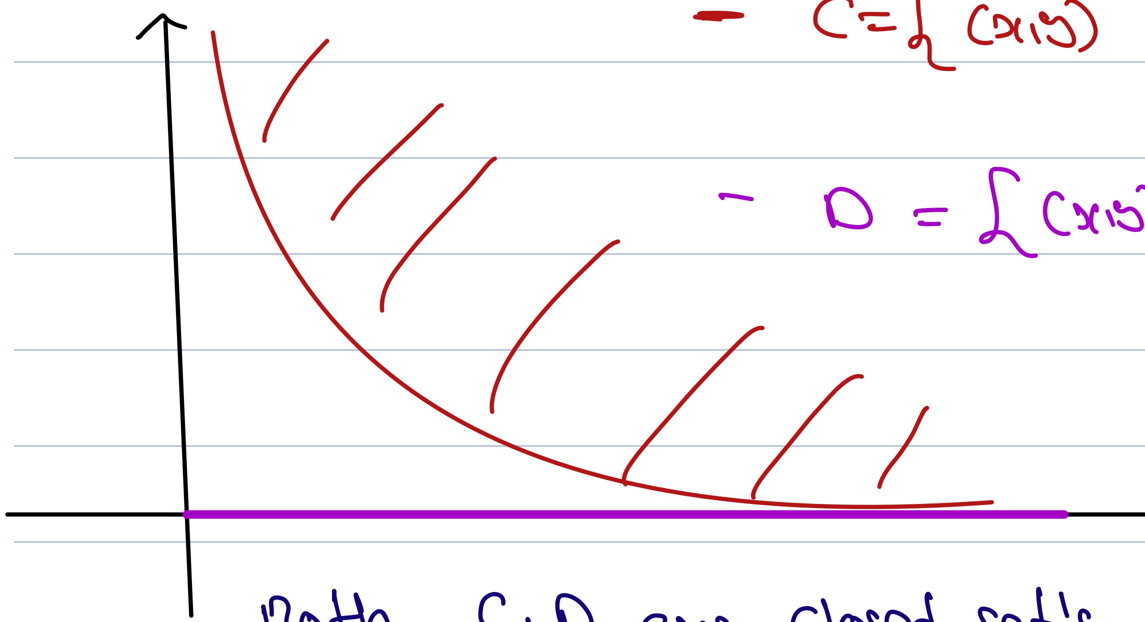
Both the set's C & D are disjoint. the only seperating hyperplane here is $a^T x = b$ (set C itself)

$\Rightarrow$ we cannot find a seperating hyperplane which is strict.

* one of these set's is open: seem's like when any of these are open set, then we may not be able to find hyperplane that strictly seperates them.

$\Rightarrow$ Suppose the two sets are convex & Both are closed, can we find a hyperplane that seperates them strictly?

NO



- $C = \{(x,y) : y \geq \frac{1}{x}, x > 0\}$
  closed

- $D = \{(x,y) : y = 0\}$

Both C, D are closed set's, Disjoint
But the only seperating hyper plane is
D (y=0), so we cannot find strict
seperating hyperplane.

* So, even if we add the Assumption, we may not be able to find a seperating hyperplane that strictly seperates them.

* turn's out if we have Boundedness on top of closedness then we can find strict seperating hyperplane.

is the inverse correct? if two set's C and D have a seperating hyperplane, then they must be disjoint?
[TRUE / FALSE] Ex: $C=\{0\}$ $D=\{0\}$ (hyperplane goes through 0)

* But inverse is correct. if atleast one of them are open

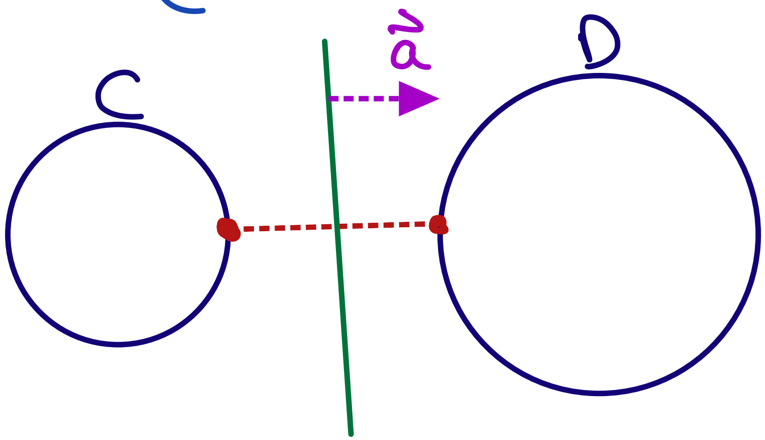
Proof of separating hyperplane theorem [simpler version]

Assumption

1. suppose $\text{dist}(C, D) = \inf\{\|x-y\| : x \in C, y \in D\} > 0$

this excludes the case where we have hyperplane and open half space becau dist = 0

2. The infimum is attained for some $c \in C$ $d \in D$ (inf does not happen at ∞)



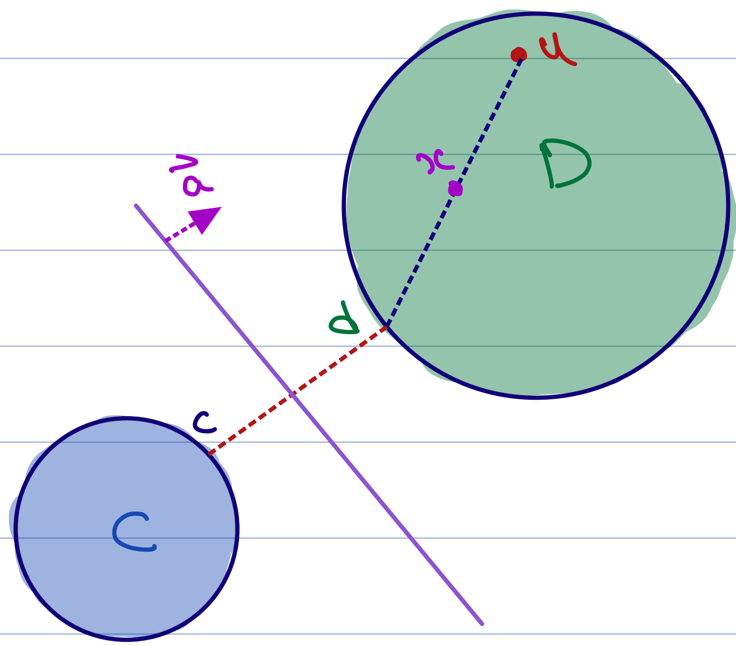
want to find a and b s.t $a^T x \geq b$ for $x \in C$

$a^T x \leq b$ for $x \in D$

Define: $a = d - c \in \mathbb{R}^n$

$$b = \frac{\|d\|_2^2 - \|c\|_2^2}{2} \in \mathbb{R} \text{ (scalar)}$$

we will show that $f(x) = a^T x - b$ is negative on C and positive on D



Proof by contradiction:

suppose $\exists u \in D$ s.t $a^T u - b \leq 0$

$$\Rightarrow f(u) \leq 0$$

$$\Rightarrow f(u) = a^T u - b$$

$$\Rightarrow (d-c)^T u - \frac{\|d\|_2^2 - \|c\|_2^2}{2} \leq 0$$

$$\|d-c\|_2^2 = (d-c)^T (d-c)$$
$$= d^T d + c^T c - 2 d^T c$$
$$= \|d\|^2 + \|c\|^2 - 2 d^T c$$

$$\Rightarrow f(u) = a^T u - b = (d-c)^T u - \frac{\|d\|_2^2 - \|c\|_2^2}{2} \leq 0$$

$$\Rightarrow (d-c)^T u + \frac{1}{2}\left(c^T c - d^T d\right) \leq 0$$

$$\Rightarrow (d-c)^T u + \frac{1}{2}\left(c^T c + d^T d - 2c^T d + 2c^T d - 2d^T d\right)$$

$$\Rightarrow (d-c)^T u + \frac{1}{2}\left(\|d-c\|_2^2 + 2(c-d)^T d\right)$$

$$\Rightarrow (d-c)^T u + \frac{1}{2}\|d-c\|_2^2 - (d-c)^T d \leq 0$$

$$\Rightarrow \underbrace{(d-c)^T (u-d)}_{<0} + \underbrace{\frac{1}{2}\|d-c\|_2^2}_{\frac{1}{2}\{\text{dist}(c,d)\} > 0} \leq 0$$

Assumed strictly +ve

$$\Rightarrow (d-c)^T (u-d) < 0$$

Let us consider $x = (1-\theta) d + \theta u \in D$

$$\Rightarrow \|x-c\|_2^2 = \|(1-\theta)d + \theta u - c\|_2^2$$

$$= \|d-c + \theta(u-d)\|_2^2$$

$$= \|d-c\|_2^2 + \underbrace{\theta^2 \|u-d\|_2^2 + \underbrace{2\theta (d-c)^T (u-d)}_{\text{strictly -ve}}}$$

$\|x-c\|_2^2 < \|d-c\|_2^2$ $\Leftarrow$ for sufficiently small value of θ we can make this summation -ve

$\Downarrow$

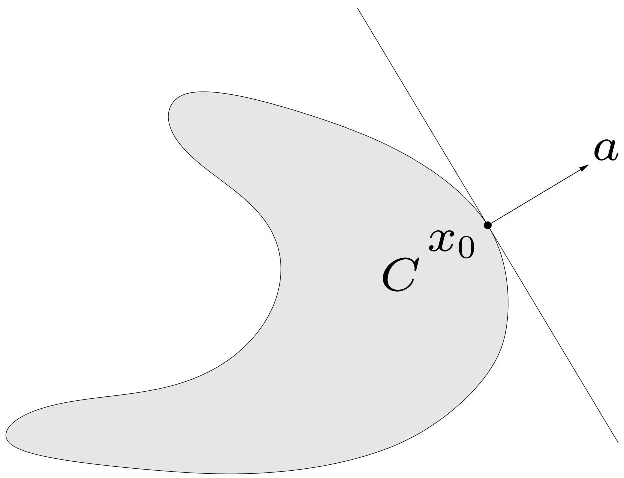
so, we could find an $x \in D$ that is closer to set C (contradiction)

Supporting hyperplane theorem

supporting hyperplane to set C at boundary point x_0:

$$\{x \mid a^T x = a^T x_0\}$$

where $a \neq 0$ and $a^T x \leq a^T x_0$ for all $x \in C$

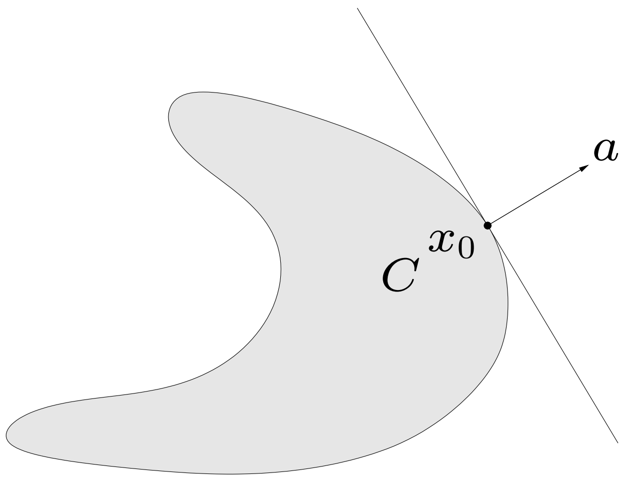


Supporting hyperplane theorem

supporting hyperplane to set C at boundary point x_0:

$$\{x \mid a^T x = a^T x_0\}$$

where $a \neq 0$ and $a^T x \leq a^T x_0$ for all $x \in C$



supporting hyperplane theorem: if C is convex, then there exists a supporting hyperplane at every boundary point of C

Boundary, interior, and closure of a set

Interior point of a set C: x is an interior point of C if there exists an open ball centered at x that is contained entirely in C.
Interior of a set (int C): The set of all interior points of C.
Open set: C is open iff $C = \textbf{int}\, C$. Ex: $C = \{x : x > 0\}$

open set may not contain limiting points of any seq^n inside the set

Ex: $x_k = \frac{1}{k}$ for any $k > 0$, $\lim_{k \to \infty} x_k = 0$ $0 \notin C$

Closure of a set (cl C): The set C plus its limiting points.

$\text{cl}\, C = \{x : x \geq 0\}$

Closed set: C is closed iff $C = \textbf{cl}\, C$.
Boundary of a set (C): $\textbf{bd}\, C = \textbf{cl}\, C \setminus \textbf{int}\, C$

Sketch of the proof of supporting hyperplane theorem

TWO CASES:-

CASE I : interior of C is non empty.

Suppose $x_0 \in \text{bd}\, C$

$D = \{x_0\}$, set D is a single Boundary points of C

$E = \{\text{int}\, C\}$, set E is interior of set C

$\Rightarrow$ Now, Both D , E are disjoint, nonempty, Convex $\Rightarrow$ apply seperating hyperplane theorem.

CASE 2 : interior of C is empty

if C is convex, empty interior $\Rightarrow$ It is hyperplane (Affine set)

i.e we don't have any curve

$\Rightarrow$ The hyperplane that contain C is itself seperating hyperplane, that seperates x_0 & rest of C.