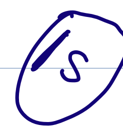
DSA3102 Lecture 2: Reading Sec 2.3, 2.4.1, 2.5, 2.6.1, 2.6.2

Announcements: 1) HW1 out (Due 07/09/2023)
2) Tutorials start next week (tut0.pdf)
3) Office hours every Tues @ 3:30pm @ S17-05-20.

If minimizing a convex function over a convex set then every local minimum is a global minimum

finite dimension euclidean space

Recap of $\boxed{\text{convex}}$ sets: A set $S \subseteq \mathbb{R}^n$ is convex if $\forall x, y \in S$, $\theta x + (1-\theta) y \in S$, $\forall \theta \in [0,1]$.



Operations that preserve convexity

union does not preserve convexity

holds for finitely many intersections

1) Intersection: $S_1, S_2 \subseteq \mathbb{R}^n$ convex $\Rightarrow S_1 \cap S_2$ is convex.

Pf: $x, y \in S_1 \cap S_2$. Then $x \in S_1$, $x \in S_2$, $y \in S_1$, $y \in S_2$.

Let $\theta \in [0,1]$. Consider

weight of convex combination

$\theta x + (1-\theta) y \in S_1 \quad \because x \in S_1, y \in S_1, S_1$ convex

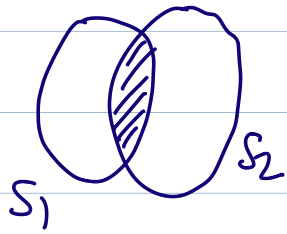
$\theta x + (1-\theta) y \in S_2 \quad \because x \in S_2, y \in S_2, S_2$ convex

$\theta x + (1-\theta) y \in S_1 \cap S_2$ (by definition since point lies in both S_1 and S_2)

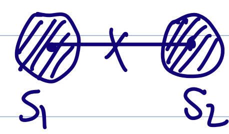
$\Rightarrow S_1 \cap S_2$ is convex ////.

collection/class

Suppose $S_\alpha, \alpha \in \mathcal{A}$ is a family of convex sets. Then $\bigcap_{\alpha \in \mathcal{A}} S_\alpha$ is convex.



Union



C.S. inequality

$|u^T x| \leq \|u\| \|x\|$

inner product

Ex: Second-order cone

$$C = \{ (x, t) \in \mathbb{R}^n \times \mathbb{R}_+ : \|x\|_2 \leq t \}.$$

x: vector; t: non-negative scalar

Fact: C is convex

Pf: The condition $\|x\|_2 \leq t \iff \forall u \in \mathbb{R}^n$ with $\|u\|_2 \leq 1$, $u^T x \leq t$

LHS: $\|x\|_2 \leq t$. Suppose $\exists u$ with $\|u\| \leq 1$ s.t. $u^T x > t$. Consider $\sup_{u : \|u\|_2 \leq 1} u^T x > t$. But the LHS of the preceding inequality is exactly $\|x\| > t$, a contradiction.

u is used as an index

Rewrite set $C = \bigcap_{u \in \mathbb{R}^n : \|u\| \leq 1} \{ (x, t) \in \mathbb{R}^n \times \mathbb{R}_+ : u^T x \leq t \}$.

$$= \bigcap_{u : \|u\| \leq 1} \underbrace{\left\{ \begin{bmatrix} x \\ t \end{bmatrix} \in \mathbb{R}^n \times \mathbb{R}_+ : \begin{bmatrix} x \\ t \end{bmatrix}^T \begin{bmatrix} u \\ -1 \end{bmatrix} \leq 0 \right\}}_{C_u}$$

$$\begin{bmatrix} x \\ t \end{bmatrix}^T \begin{bmatrix} u \\ -1 \end{bmatrix} \leq 0 \quad \Leftrightarrow \quad x^T u - t \leq 0$$

each 1 of this set is characterised by an set C_u (c indexed by vector u)

Each C_u is a $\boxed{\text{halfspace}}$ in $\mathbb{R}^{n+1}$.

Halfspaces are convex $\Rightarrow$ C is also convex.

(C is an intersection of many halfspaces)

Ex: $S_+^n = \{ X \in S^n : X \succeq 0 \}$ ← symmetric

↑ PSD.

Claim: S_+^n is convex   non-strict inequality

Pf:

for a fixed z, look at all symmetric matrices, then run over all possible z

$$S_+^n = \bigcap_{z \in \mathbb{R}^n} \underbrace{\{ X \in S^n : z^T X z \geq 0 \}}_{S_z \text{ (indexed by } z)}$$

subset of set of symmetric matrices

$S_z \subset S^n$

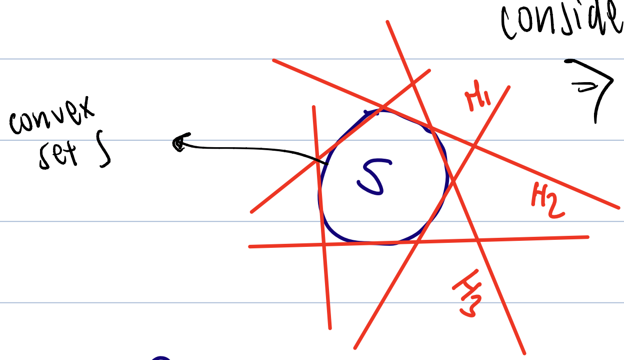
$S_z = \{ X \in S^n : \underbrace{z^T X z}_{\text{linear in } X} \geq 0 \}$ (number) is a halfspace hence convex

function takes in a matrix and gives a number

Since S_+^n is an infinite intersection of convex sets, S_+^n is convex.

Rmk: Every closed convex set S is the intersection of all halfspaces that contain it

consider all tangents



⇒ this convex set is the intersection of halfspaces that contain it

$S = \bigcap \{ H : H \text{ halfspace}, S \subset H \}$

convex sets can be decomposed into intersections of halfspaces

---

Affine function

$f: \mathbb{R}^n \to \mathbb{R}^m$ is affine if it is the sum of a linear function plus a constant vector (scalar also can)

$$\underset{\uparrow \, x \in \mathbb{R}^n}{f(x)} = \underbrace{Ax + b}_{\in \mathbb{R}^m} \qquad A \in \mathbb{R}^{m \times n}, \; b \in \mathbb{R}^m.$$

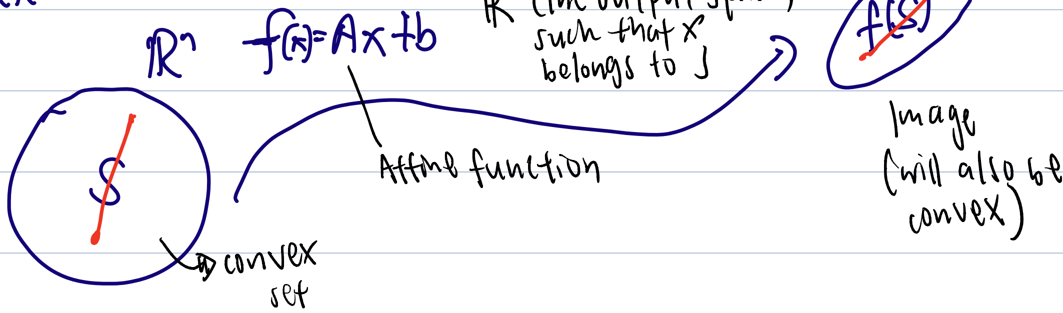
Affine function

Fact: Suppose $S \subset \mathbb{R}^n$ is convex, $f: \mathbb{R}^n \to \mathbb{R}^m$ is affine. Image of S under f

$$f(S) = \{ f(x) \in \mathbb{R}^m : x \in S \}$$

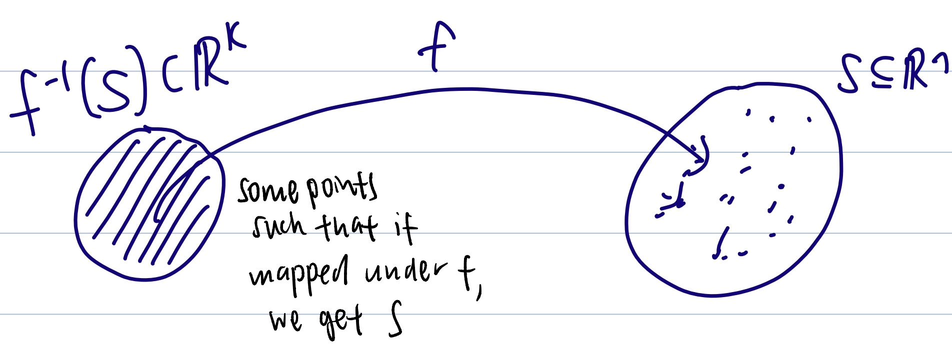
↳ set of all possible points in $\mathbb{R}^n$ (the output space) such that x belongs to S

is convex



Fact: If $f: \mathbb{R}^k \to \mathbb{R}^n$ is affine, the inverse image of S under f

$$f^{-1}(S) = \{ x \in \mathbb{R}^k : f(x) \in S \}$$



is convex

Ex: If S is convex, so are $aS = \{ax : x \in S\}$ (scale) and $x + S = \{x + y : y \in S\}$.

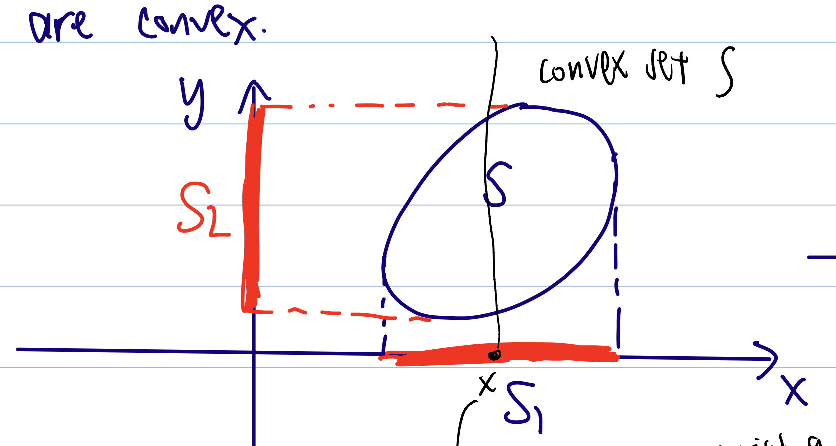
x-section / y-section

Ex: The <u>projection</u> of a convex set onto some of its coordinates is convex.

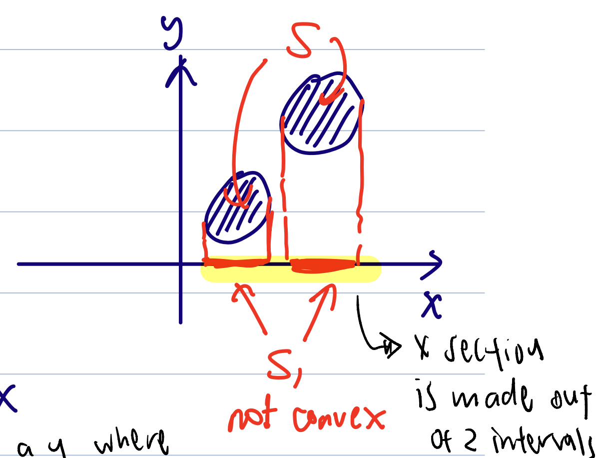
If $S \subseteq \mathbb{R}^n \times \mathbb{R}^m$ is convex

$$S_1 = \{x \in \mathbb{R}^n : (x, y) \in S \text{ for some } y \in \mathbb{R}^m\}$$
$$S_2 = \{y \in \mathbb{R}^m : (x, y) \in S \text{ for some } x \in \mathbb{R}^n\}$$

are convex.



belong to x section (exist a y where (x, y) belong inside convex set S)



Ex: The <u>sum</u> of two convex sets S_1 and S_2 is <u>convex</u>.

$$S_1 + S_2 = \{x + y : x \in S_1, y \in S_2\}.$$ newly created set is convex

Pf: Consider $S_1 \times S_2 = \{(x, y) : x \in S_1, y \in S_2\}$ (cartesian product). $S_1 \times S_2$ is convex if S_1 and S_2 are both convex.

Consider the affine f^n $f(x, y) = x + y$. Then $S_1 + S_2$ is the image of $\underbrace{S_1 \times S_2}_{\text{convex}}$ under $\underbrace{f}_{\text{affine}}$. $\Rightarrow$ $S_1 + S_2$ is convex. ////.

$$\begin{bmatrix} x \\ y \end{bmatrix}^T \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Partial sum: $S_1, S_2 \subseteq \mathbb{R}^n \times \mathbb{R}^m$

$$S = \{ (x, y_1+y_2) : (x,y_1) \in S_1, (x,y_2) \in S_2 \}$$

Fact: If S_1 and S_2 are convex, so is S.

Rmk: If $n=0$, this is set addition $S = S_1 + S_2$ convex

first part is not present (only considering y_1+y_2)

If $m=0$, ... $S = S_1 \cap S_2$ convex
(no y component)

Hyperbolic Cone

resembles $\|x\| \le t$ (2nd order cone)

$P \in S^n_+$, $c \in \mathbb{R}^n$
(positive semi definite)

P expressed as its eigen decomposition
$U \boxed{D} U^T$
diagonal matrix
(eigenvalues on its diagonals are non-negative)
↳ can take square root

$$C = \{ x \in \mathbb{R}^n : x^T P x \le (c^T x)^2, \; c^T x \ge 0 \}$$

↳ quadratic form must be non-negative

Claim: C is convex.

$\boxed{\text{2nd order cone}}$

Consider $C' = \{ (z,t) \in \mathbb{R}^n \times \mathbb{R} : \underbrace{\|z\|^2}_{\text{(non -ve)}} \le t^2, \; t \ge 0 \}$ is convex.

Consider the function $f(x) = \begin{bmatrix} P^{1/2} x \\ c^T x \end{bmatrix} \in \mathbb{R}^{n+1}$, $x \in \mathbb{R}^n$

$f : \mathbb{R}^n \to \mathbb{R}^{n+1}$

It is linear!

inverse image of C' under linear f^n f.

Claim: $C = \{ x \in \mathbb{R}^n : f(x) \in C' \}$. ← convex.

f ↑ linear, C' ↑ convex

$$= \{ x : \begin{bmatrix} P^{1/2} x \\ c^T x \end{bmatrix} \in C' \}$$

$$V = \{x : \|P^{1/2}x\|^2 \le (c^T x)^2,\ c^T x \ge 0\}$$
$$= \{x : x^T P x \le (c^T x)^2,\ c^T x \ge 0\}$$

Perspective Transform

looks a little non-linear but still preserves convexity

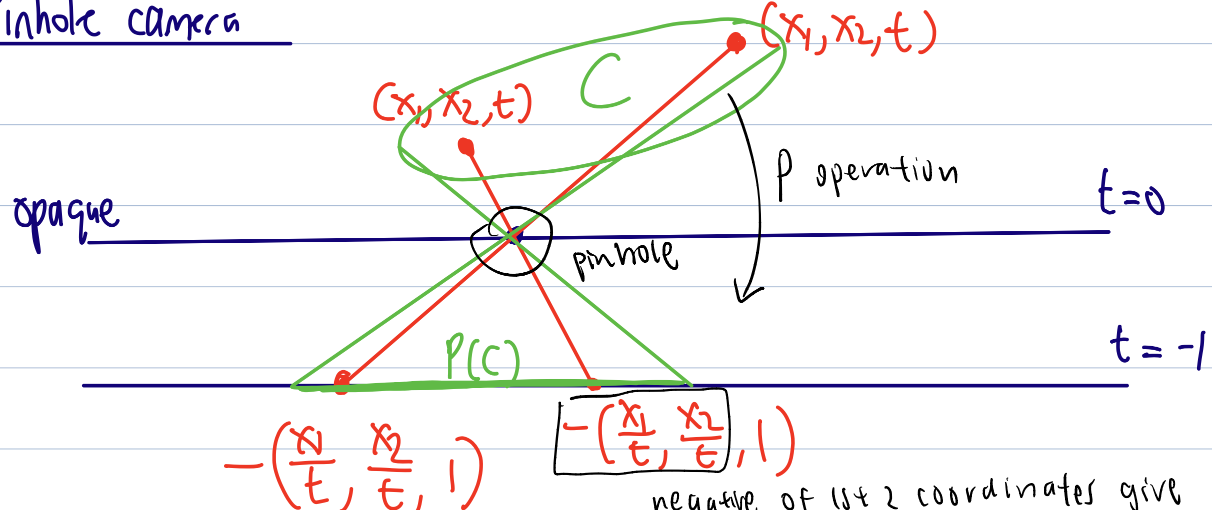
Perspective function $P: \mathbb{R}^{n+1} \to \mathbb{R}^n$ $\quad$ $\text{dom } P = \mathbb{R}^n \times \mathbb{R}_{++}$

$\mathbb{R}_+ = [0, \infty)$, $\quad \mathbb{R}_{++} = (0, \infty)$

$$P(z, t) = \frac{z}{t}$$

z: vector in $\mathbb{R}^n$ $\quad$ t: positive number

Pinhole camera



t must be +ve to pass through the function

Fact 1: If $C \subseteq \text{dom } P \subseteq \mathbb{R}^{n+1}$ is convex. Then

$P(C) = \{P(x,t) : (x,t) \in C\} \subset \mathbb{R}^n$ convex. $\Big\}$ closed under perspective transform

image of C under the perspective transform $\quad P(x,t) = \frac{x}{t}$

Fact 2: If $C \subseteq \mathbb{R}^n$ is convex, then

$P^{-1}(C) = \{(x,t) \in \mathbb{R}^{n+1} : \frac{x}{t} \in C,\ t > 0\}$ is convex.

inverse image

Proof of Fact 1: NTS: $\forall \theta \in [0,1]$, $(x,t), (y,s) \in C$

$$\theta \frac{x}{t} + (1-\theta)\frac{y}{s} \in P(C)$$

$x, y \in \mathbb{R}^n$, $s, t \in \mathbb{R}_{++}$.

$$\theta \frac{x}{t} + (1-\theta)\frac{y}{s}$$

$$= \frac{\theta x s + (1-\theta) y t}{st}$$

$$= \frac{\dfrac{\theta x s + (1-\theta) y t}{\theta s + (1-\theta) t}}{\dfrac{st}{\theta s + (1-\theta) t}}$$

convex combination inside $\mathbb{R}_{++}$

$$= \frac{\dfrac{\theta s}{\theta s + (1-\theta)t} x + \dfrac{(1-\theta)t}{\theta s + (1-\theta)t} y}{\dfrac{\theta s}{\theta s + (1-\theta)t} t + \dfrac{(1-\theta)t}{\theta s + (1-\theta)t} s}$$

$$= \frac{\mu x + (1-\mu) y}{\mu t + (1-\mu) s}$$

where $\mu = \dfrac{\theta s}{\theta s + (1-\theta) t}$.

convex combination weights all add up to 1 hence entire thing belongs to C

Since $C \subseteq \mathbb{R}^{n+1}$ is convex, $\mu \underbrace{\begin{bmatrix} x \\ t \end{bmatrix}}_{\in C} + (1-\mu) \underbrace{\begin{bmatrix} y \\ s \end{bmatrix}}_{\in C} \in C$

$P(C) = \{ \frac{x}{t} : (x,t) \in C \}$

Hence, by the definition of $P(C)$ $\quad \dfrac{\mu x + (1-\mu) y}{\mu t + (1-\mu) s} \in P(C)$

Therefore $P(C)$ is convex. ////.

$P(x,t) = \frac{x}{t}$ $\quad C \subset \operatorname{dom} P$ $\quad$ C contains pts of

$P: \mathbb{R}^n \times \mathbb{R}_{++} \to \mathbb{R}^n$ $\quad = \mathbb{R}^n \times \mathbb{R}_{++}$ $\quad$ the form (x,t)

$x \in \mathbb{R}^n, t > 0$

$$P(C) = \{ P(x,t) : (x,t) \in C \}$$
$$= \{ \frac{x}{t} \in \mathbb{R}^n : (x,t) \in C \}$$

Linear fractional functions: Composition of perspective with affine.

$g: \mathbb{R}^n \to \mathbb{R}^{m+1}$ $\quad g(x) = \begin{bmatrix} A \\ c^T \end{bmatrix} x + \begin{bmatrix} b \\ d \end{bmatrix}$ affine

linear function $\in \mathbb{R}^n$ $\quad$ constant vector

$A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$ $\quad d \in \mathbb{R}^1, c \in \mathbb{R}^n$

(scalar)

bottom part

$f: \mathbb{R}^n \to \mathbb{R}^m, \quad f = P \circ g$

$f(x) = P(g(x))$ $\quad \operatorname{dom} f = \{ x : c^T x + d > 0 \}$

is called linear-fractional function

want to divide by bottom component (so cannot be 0)

↳ linear fractional function

$$f(x) = \frac{\overbrace{Ax+b}^{\in \mathbb{R}^m}}{\underbrace{c^T x + d}_{\text{scalar} > 0}}$$

$\frac{\text{vector}}{\text{number}}$ = vector in $\mathbb{R}^m$

Ex: Conditional Probabilities.

$U \in \{1, \dots, n\}$ $\quad$ $V \in \{1, \dots, m\}$

$p_{ij} = \Pr(U=i, V=j)$ $\quad$ $i \in [n],\ j \in [m]$

Conditional Prob $\Pr(U=i \mid V=j) = \dfrac{\Pr(U=i, V=j)}{\Pr(V=j)}$

$$= \frac{\Pr(U=i, V=j)}{\sum_{k=1}^{n} \Pr(U=k, V=j)}$$

joint probability is a matrix of numbers that is non-negative | adds up to 1

$$p_{i|j} = \frac{p_{ij}}{\sum_k p_{kj}}$$

linear function of matrix

gives a certain number

Treat $\{p_{ij} : i \in [n], j \in [m]\}$ as a matrix

The cond. prob. $\{p_{i|j} : i \in [n], j \in [m]\}$ can be obtained as a linear fractional map from $\{p_{ij}\}$

So C, a convex set of joint dist. on support $[n] \times [m]$, then the associated set of conditional probabilities is convex.

↳ passing through linear fractional transform (which is composition of perspective with affine function ⇒ preserves convexity)

Proper cones & generalized inequalities

$K \subseteq \mathbb{R}^n$: cone $\quad \forall x \in K,\ \lambda x \in K \quad \forall \lambda \geq 0.$

K is a <u>proper cone</u> if

i) convex ii) closed iii) solid (non-empty interior)

iv) pointed (no empty boundaries)

induces a

<u>Partial order</u>

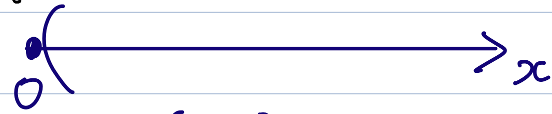
according to proper cone K

$x \preceq_K y \iff y - x \in K.$

$x \prec_K y \iff y - x \in \text{int}(K)$

non-negative number is a cone

Ex: $K = \mathbb{R}_+ = [0, \infty)$

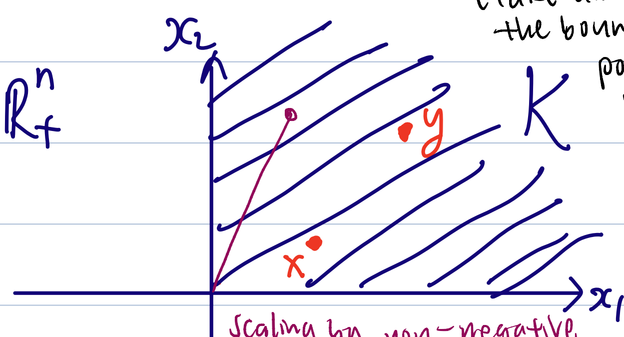


$3 \prec_K 5 \Rightarrow 5 - 3 = 2 \in \text{int}(K),\ 2 \in (0, \infty)$

$3 \not\prec_K 3 \Rightarrow 3 - 3 = 0 \in K$

$\text{int}(\mathbb{R}_+) = \mathbb{R}_{++}$ (take away the boundary point which is 0)

Ex: Non-negative <u>orthant</u> $\mathbb{R}^n_+$



scaling by non-negative number $\Rightarrow$ still inside the cone

$x \preceq_K y \qquad x \preceq_{\mathbb{R}^n_+} y$

iff $x_i \le y_i \ \forall i = 1, \dots, n$

$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},\ y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$

$\text{int}(\mathbb{R}^n_+) = \mathbb{R}^n_{++}$

length n vector such that each of the entry is positive

$y - x \in \mathbb{R}^n_+$

$x = \begin{pmatrix} 2 \\ 1 \end{pmatrix},\ y = \begin{pmatrix} 5 \\ 1 \end{pmatrix},\quad y - x = \begin{pmatrix} 4 \\ 0 \end{pmatrix} \in \mathbb{R}^2_+ = K,\quad x \preceq_K y$

$\cancel{x \prec_K y} \qquad y - x = \begin{pmatrix} 4 \\ 0 \end{pmatrix} \notin \text{int}(K)$

(zero is not positive)

$x = \begin{pmatrix} 2 \\ 1 \end{pmatrix},\ y = \begin{pmatrix} 1 \\ 2 \end{pmatrix},\quad x \not\preceq y \qquad y \not\preceq x$

Not all pairs of vectors can be compared positive definite matrices

Ex: PSD cone $S_+^n = K$ $\quad \text{int}(K) = S_{++}^n$

$S_+^n = K$ is a proper cone (induces an ordering amongst the semi-definite matrices)

$X \preccurlyeq_K Y \iff Y - X \in K \iff Y - X$ is PSD.

$X \prec_K Y \iff Y - X \in \text{int}(K) \iff Y - X$ is positive def.

symmetric matrix $X = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$, $\quad Y = \begin{pmatrix} 10 & 8 \\ 8 & 10 \end{pmatrix}$

$X \preccurlyeq_K Y \qquad Y - X = \begin{pmatrix} 9 & 6 \\ 6 & 9 \end{pmatrix} \succcurlyeq 0$

check positive semi-definite using determinant

$\det(Y - X) = 9^2 - 6^2 > 0.$

---

1) $x \preccurlyeq_K y$, $u \preccurlyeq_K v \Rightarrow x + u \preccurlyeq_K y + v$
2) $x \preccurlyeq_K y$, $y \preccurlyeq_K z \Rightarrow x \preccurlyeq_K z$
3) $x \preccurlyeq_K y$, $\alpha x \preccurlyeq_K \alpha y$.

$\vdots$

Thm: Separating Hyperplane Thm.

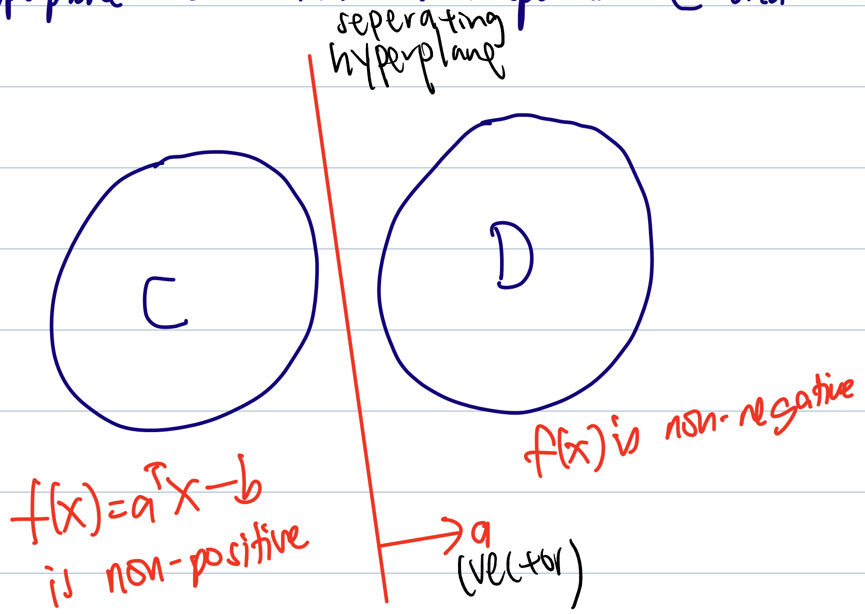
Suppose C & D are non-empty disjoint convex sets ($C \cap D = \emptyset$)

Then $\exists\, a \neq 0$, b s.t. $a^T x \leq b \quad \forall x \in C$

$a^T x \geq b \quad \forall x \in D$

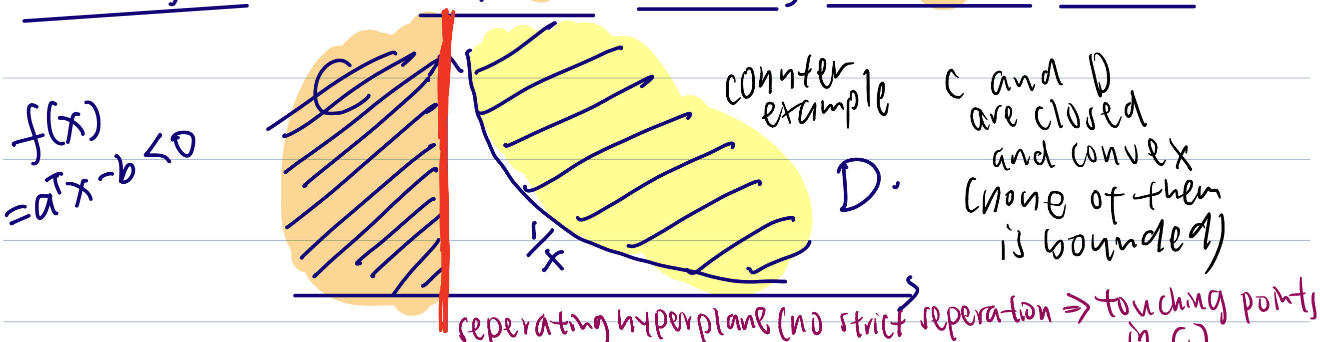
The affine function $f(x) = a^T x - b$ is non-positive on C & non-negative on D.

The hyperplane $\{x : a^T x = b\}$ separates C and D.



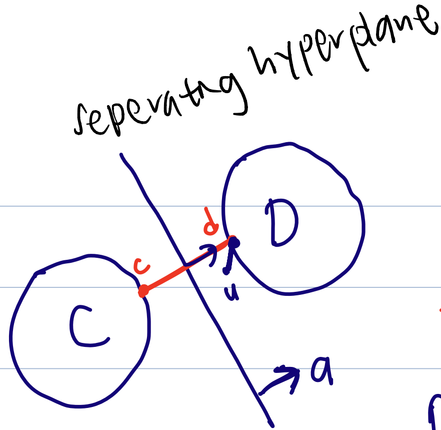
Rmk: If both sets are closed and at least one of them is bounded, then separation can be strict, i.e., inequality

$\exists\, a \neq 0, b$ s.t. $a^T x < b \quad \forall x \in C$, $a^T x > b$, $\forall x \in D$.



HW: Use strict sep. hyperplane thm to prove linear programming duality.

Pf: (Sep. hyperplane thm) C, D convex disjoint



$\text{dist}(C, D) = \inf\{\|c-d\| : c \in C, d \in D\} > 0.$

Say points $c \in C$, $d \in D$ achieve the inf

points on set C and D that are closest to each other

Define $a = d - c$, $b = \dfrac{\|d\|_2^2 - \|c\|_2^2}{2}$

This is the desired affine fn.

$$f(x) = a^T x - b$$
$$= (d-c)^T x - \left(\frac{\|d\|_2^2 - \|c\|_2^2}{2}\right)$$
$$= (d-c)^T\left(x - \tfrac{1}{2}(d+c)\right)$$

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

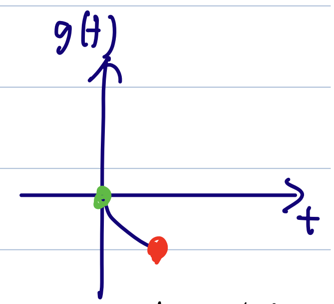
non-strict version (symmetric statements)

✓ 1) f is non-negative on D $\quad f(x) = a^T x - b \geq 0 \quad \forall x \in D.$

2) f is non-positive on C.

Suppose, to the contrary, $\exists u \in D$ s.t. $f(u) < 0$.

$$f(u) = (d-c)^T\left(u - \tfrac{1}{2}(d+c)\right)$$
$$= (d-c)^T\left(u - d + \tfrac{d}{2} - \tfrac{c}{2}\right)$$
$$= \underbrace{(d-c)^T(u-d)}_{<0} + \underbrace{\tfrac{1}{2}\|d-c\|^2}_{\geq 0 \text{ (norm)}} < 0$$



(derivative of a certain function less than 0)

In summary $(d-c)^T(u-d) < 0$

Consider $\displaystyle \frac{d}{dt} \|\overbrace{\underbrace{d + t(u-d)}_{\text{convex combi of } d \text{ and } u}}^{g(t)} - c\|_2^2 \Big|_{t=0} = 2(d-c)^T(u-d) < 0$

For small enough $t > 0$, red point green point

$$\|d+t(u-d)-c\|_2 < \|d-c\|$$

The point $d+t(u-d)$ is closer to c than d.

This, however, contradicts the optimality of d.

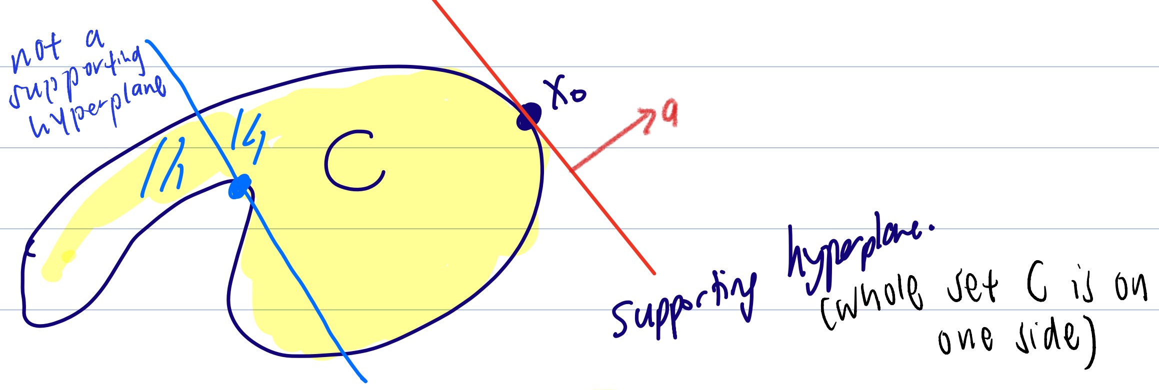
that C and D are closest to each other

$C \subseteq \mathbb{R}^d$ closed & bounded $\Rightarrow$ compact.

Supporting hyperplane

boundary of C (closure − interior of C)

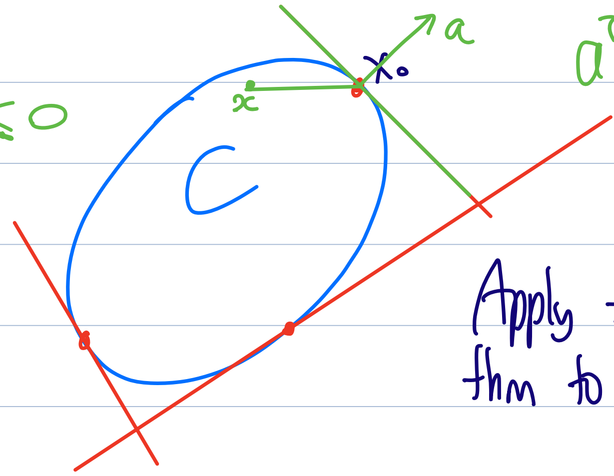
$C \subseteq \mathbb{R}^n$ $\quad x_0 \in Bd(C) = Cl(C) \setminus Int(C)$.



If $a \neq 0$ satisfies $a^T x \leq a^T x_0$ $\forall x \in C$, then the set $\{x : a^T x = a^T x_0\}$ is called a separating hyperplane of C at x_0.

Thm: For every non-empty convex set C and any $x_0 \in Bd(C)$, $\exists$ a supporting hyperplane of C at x_0.

$a^T(x - x_0) \leq 0$



$a^T x \leq a^T x_0.$
$\forall x \in C$

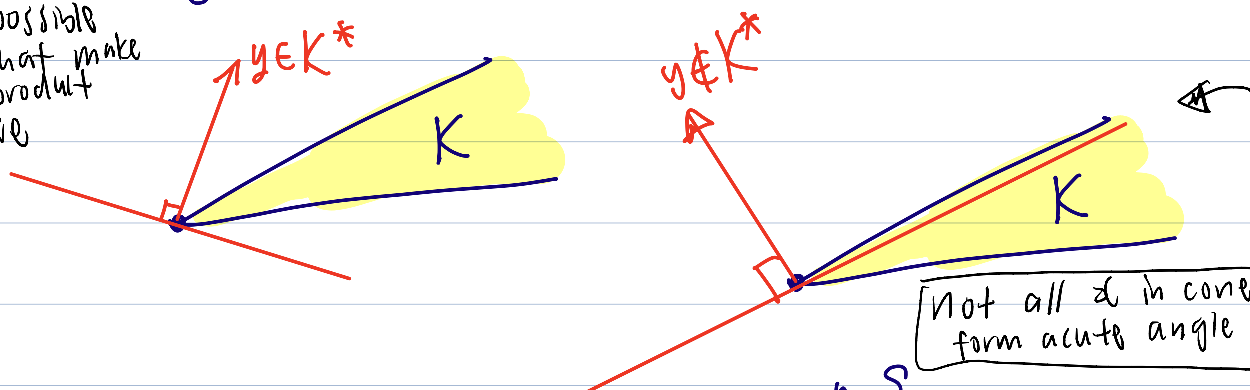
Apply the sep. hyperplane thm to $\{x_0\}$ and int(C)

Pf: Application of separating hyperplane thm.

Dual cones: K: cone. $\forall x \in K$, $\lambda x \in K$ $\forall \lambda \geq 0$.

$K^* = \{y : \underbrace{x^T y \geq 0}_{\text{acute angle between } x \text{ \& } y.} \ \forall x \in K\}$. dual cone of K.

Find all possible vectors that make inner product positive



Not all x in cone form acute angle with y

Ex: $K = \mathbb{R}^n_+$: non-negative orthant $\mathbb{R}^n_+ = \{x \in \mathbb{R}^n : x_i \geq 0 \ \forall i \in [n]\}$.

each component is non-negative

Claim: $K^* = K$ self-dual.

dual cone = original cone

$y^T x \geq 0 \ \forall x \geq 0 \iff y \geq 0$ $(y_1, \ldots, y_n)$

Clearly $\Leftarrow$ holds true. (If y is non-negative, inner product will sure be non-negative)

$\Rightarrow$ Suppose, to the contrary, $\underbrace{\exists i \in [n] \text{ s.t. } y_i < 0}_{\text{suppose one component in the vector is negative}}$.

Then set $x = e_i \geq 0$. Then $y^T x = y_i < 0$. This contradicts that $y^T x \geq 0$.

$$\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \leftarrow \text{ith location.}$$

$\Rightarrow y \geq 0$ ////

$K^* = K$.

Ex: $(S_+^n)^* = S_+^n$

positive semi-definite cone is also self-dual