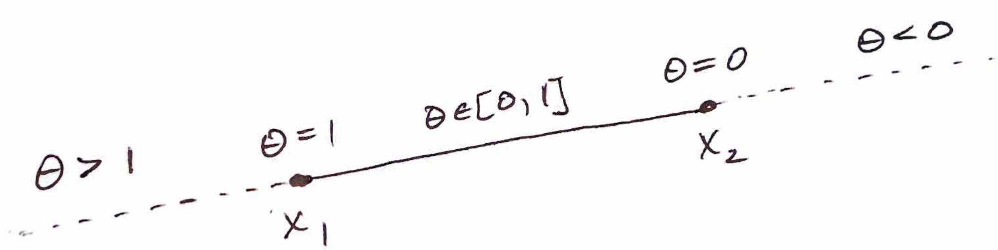
Lecture 5 : Convex Sets

goals:
- intro to convex sets
- examples of convex sets
- operations that preserve convexity
- generalized inequalities
- separating + supporting hyperplane theorems

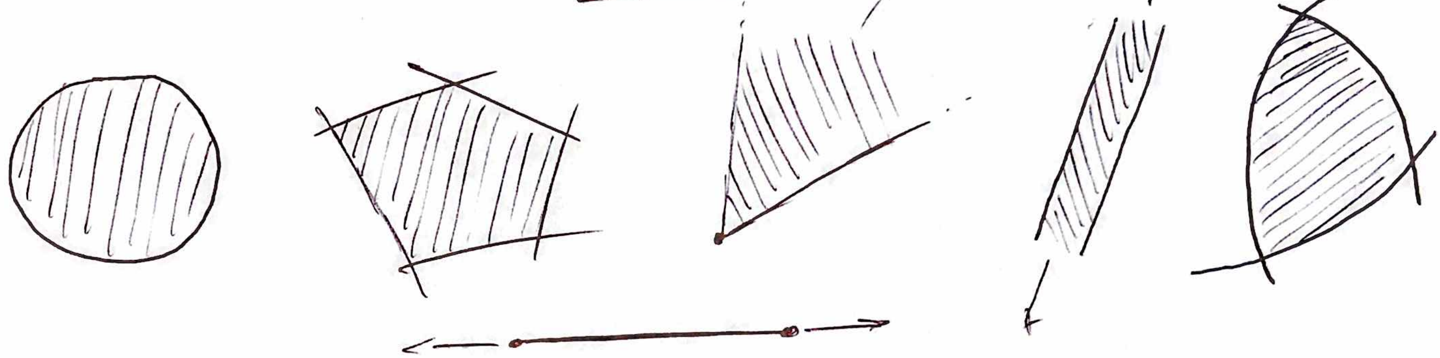
Definition: A set $\mathcal{X} \subseteq \mathbb{R}^n$ is called convex if for any two points $x_1, x_2 \in \mathcal{X}$, the convex combination of x_1 and x_2 lies in $\mathcal{X}$, i.e.

$$\theta x_1 + (1-\theta) x_2 \in \mathcal{X} \quad \forall \theta \in [0,1]$$

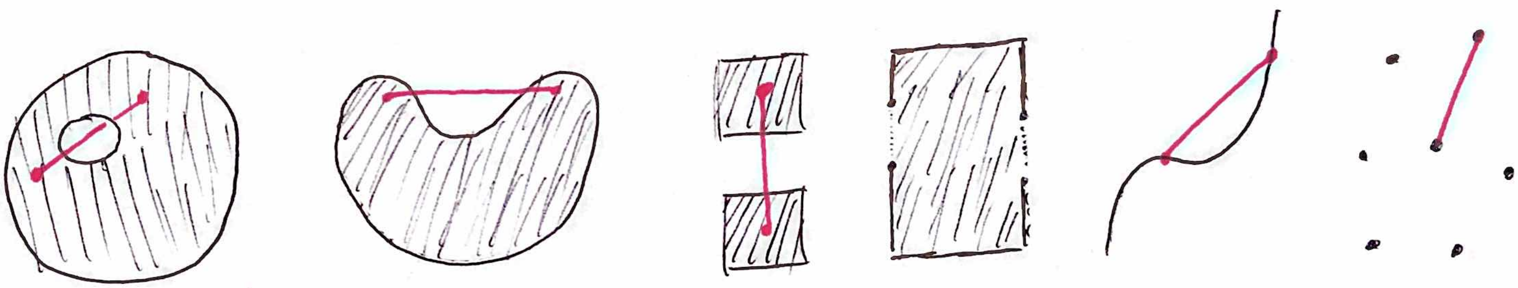
- interpretation: all line segments starting + ending in $\mathcal{X}$ stay within $\mathcal{X}$



Convex:



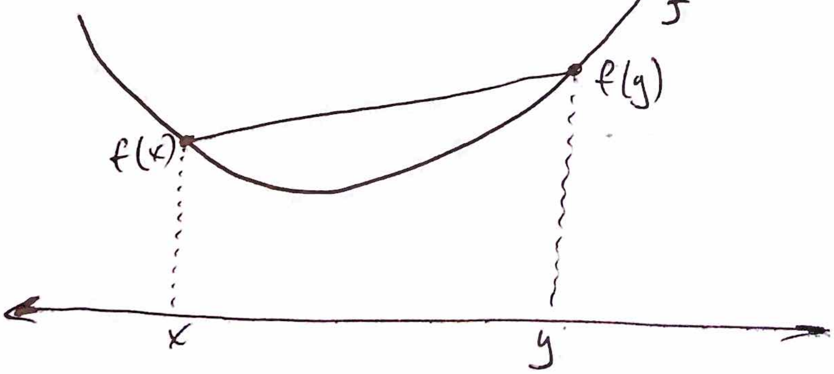
Non-Convex:



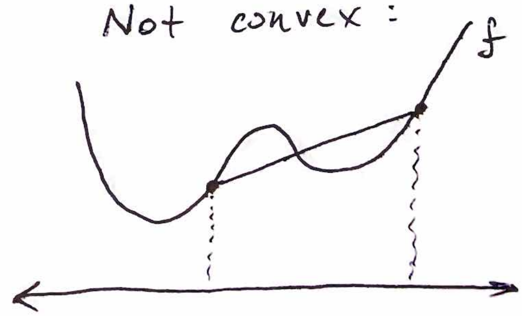
Definition: A function $f: \mathbb{R}^n \to \mathbb{R}$ is called convex if its domain $\mathrm{dom}(f)$ is a convex set and for all $x, y \in \mathrm{dom}\, f$ we have

$$f(\theta x + (1-\theta) y) \leq \theta f(x) + (1-\theta) f(y)$$

$$\forall \theta \in [0, 1]$$



Not convex:



- Theory of convex sets and convex functions connects via the *epigraph*: $\{(x,t) \in \mathbb{R}^{n+1} \mid x \in \mathrm{dom}(f), f(x) \le t\}$
  - allows switching back and forth between *geometric* and *analytical* approaches

---

Theorem A function $f: \mathbb{R}^n \to \mathbb{R}$ is convex iff its epigraph $\{(x,t) \in \mathbb{R}^{n+1} \mid x \in \mathrm{dom}(f), f(x) \le t\}$ is convex (as a set).

Proof: ($\Rightarrow$) Suppose $\mathrm{epi}(f)$ is not convex. Then $\exists\, (x, t_x), (y, t_y), \theta \in [0,1]$ such that $f(x) \le t_x$ and $f(y) \le t_y$ and

$$(\theta x + (1-\theta) y,\ \theta t_x + (1-\theta) t_y) \notin \mathrm{epi}(f).$$

This implies

$$f(\theta x + (1-\theta) y) > \theta t_x + (1-\theta) t_y \ge \theta f(x) + (1-\theta) f(y)$$

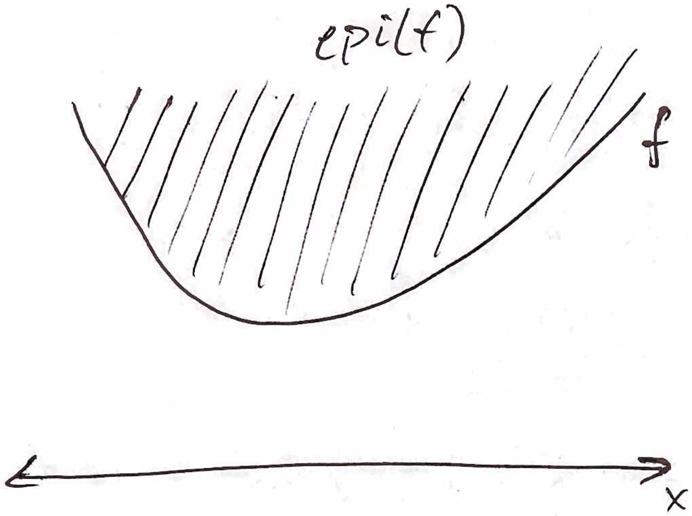
($\Leftarrow$) Suppose f is not convex. Then $\exists\, x, y \in \mathrm{dom}(f)$, $\theta \in [0,1]$ such that

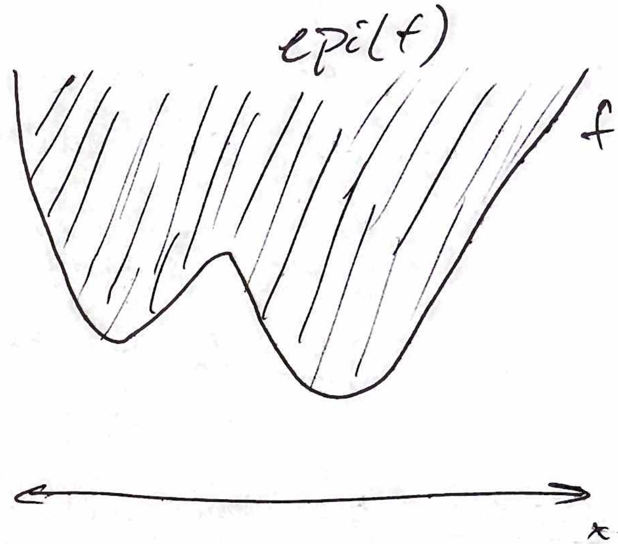
$$f(\theta x + (1-\theta) y) > \theta f(x) + (1-\theta) f(y).$$

Consider $(x, f(x)), (y, f(y)) \in \mathrm{epi}(f)$. Then

$$(\theta x + (1-\theta) y,\ \theta f(x) + (1-\theta) f(y)) \notin \mathrm{epi}(f)$$

$\blacksquare$





Theorem: Consider an optimization problem

$$\begin{aligned} &\text{minimize} && f(x) \\ &\text{subject to} && x \in \mathcal{X} \end{aligned}$$

where f is a convex function and $\mathcal{X}$ is a convex set. Then any local minimum is also a global minimum.

Proof: Let $\bar{x}$ be a local minimum, i.e. $\bar{x} \in \mathcal{X}$ and $\exists \varepsilon > 0$ s.t. $f(\bar{x}) \le f(x)$ $\forall B(\bar{x}, \varepsilon)$. To obtain a contradiction, suppose $\exists z \in \mathcal{X}$ with $f(z) < f(\bar{x})$.

By convexity of $\mathcal{X}$ we have

$$\theta \bar{x} + (1-\theta) z \in \mathcal{X} \qquad \forall \theta \in [0,1]$$

By convexity of f we have

$$\begin{aligned} f(\theta \bar{x} + (1-\theta) z) &\le \theta f(\bar{x}) + (1-\theta) f(z) \quad \forall \theta \in [0,1] \\ &< \theta f(\bar{x}) + (1-\theta) f(\bar{x}) = f(\bar{x}) \end{aligned}$$

But as $\theta \to 1$, $(\theta \bar{x} + (1-\theta) z) \to \bar{x}$, and the inequality above contradicts the local optimality of $\bar{x}$.

$\blacksquare$

- A point $\theta_1 x_1 + \cdots + \theta_k x_k$ where $\theta_1 + \cdots + \theta_k = 1$ and $\theta_i \geq 0$, $i = 1, \ldots, k$ is called a convex combination of $x_1, \ldots, x_k$

- The convex hull of a set $C \subseteq \mathbb{R}^n$, denoted $\mathrm{conv}(C)$ is the set of all convex combinations of points in C:

$$\mathrm{conv}(C) = \{\theta_1 x_1 + \cdots + \theta_k x_k \mid x_i \in C, \theta_i \geq 0, \textstyle\sum \theta_i = 1\}$$

  - see BV Fig 2.3

- A set $C \subseteq \mathbb{R}^n$ is affine if the line through any two distict points in C lies in C, i.e.

$$\underbrace{\theta x_1 + (1-\theta) x_2}_{\text{linear combination}} \in C \quad \forall x_1, x_2 \in C, \theta \in \mathbb{R}$$

Every affine set can be represented as the solution set of a system of linear equations

$$C = \{x \in \mathbb{R}^n \mid Ax = b\} \quad A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$$

If $b = 0$, the affine set is a subspace

- Clearly, all affine sets are convex

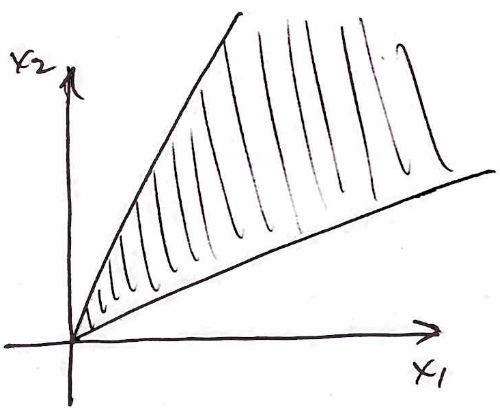
Cones

- A set $C \subseteq \mathbb{R}^n$ is called a cone if

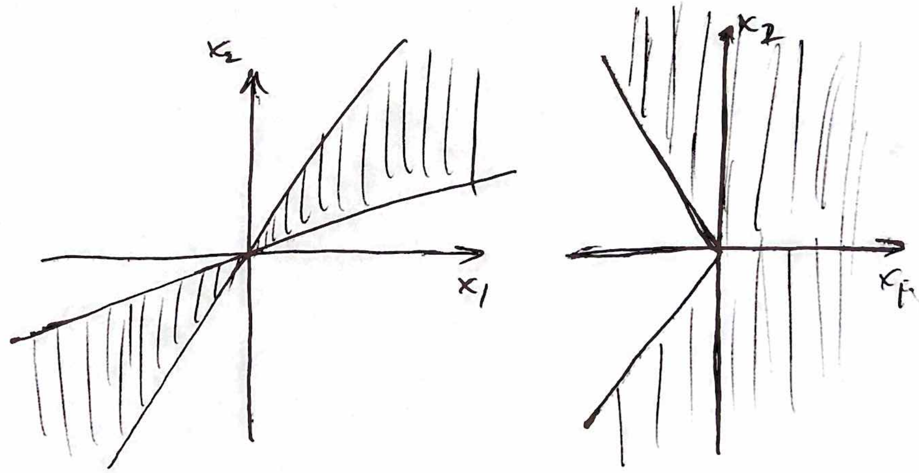
$$\theta x \in C \quad \forall x \in C \text{ and } \theta \geq 0$$

- A set $C \subseteq \mathbb{R}^n$ is a convex cone if it's convex and a cone, or equivalently if

$$\theta_1 x_1 + \theta_2 x_2 \in C \quad \forall x_1, x_2 \in C,\ \theta_1, \theta_2 \geq 0$$



a convex cone



non-convex cones

- A point $\theta_1 x_1 + \cdots + \theta_k x_k$ with $\theta_i \geq 0 \ \forall i$ is called a conic combination of $x_1, \ldots, x_k$

- The conic hull of a set C is the set of all conic combinations of points in C:

$$\{\theta_1 x_1 + \cdots + \theta_k x_k \mid x_i \in C, \theta_i \geq 0\}$$

- see BV Fig 2.5

Hyperplanes and halfspaces

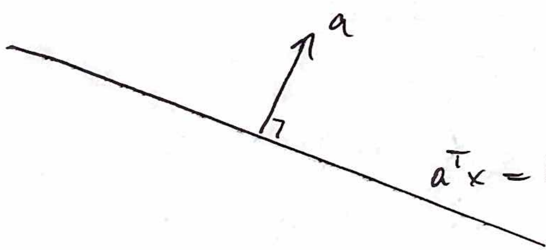
- A <u>hyperplane</u> is a set of the form

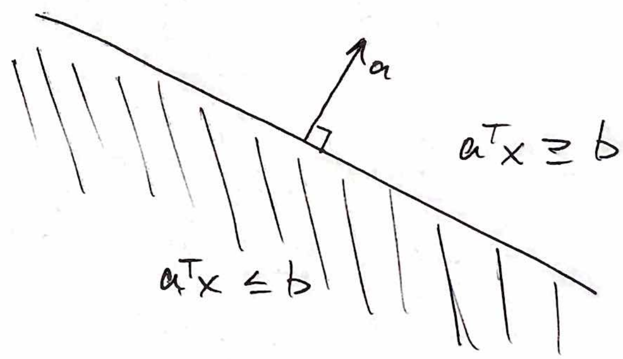
$$\{x \in \mathbb{R}^n \mid a^T x = b\}$$

where $a \neq 0$ defines the normal vector to the hyperplane

- A hyperplane divides $\mathbb{R}^n$ into two <u>halfspaces</u>.
  A <u>halfspace</u> is a set of the form

$$\{x \in \mathbb{R}^n \mid a^T x \leq b\}$$





- all hyperplanes and halfspaces are convex

Polyhedra and Polytopes

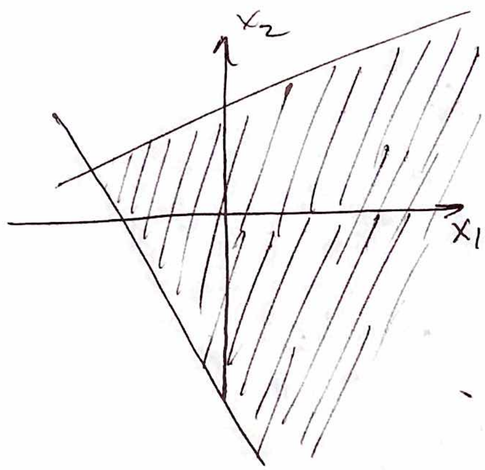
- A <u>polyhedron</u> is the solution set of a <u>finite</u> number of closed halfspaces and hyperplanes

$$P = \{x \in \mathbb{R}^n \mid a_j^T x \leq b_j,\ j=1,\dots,m,\ c_j^T x = d_j,\ j=1,\dots,p\}$$

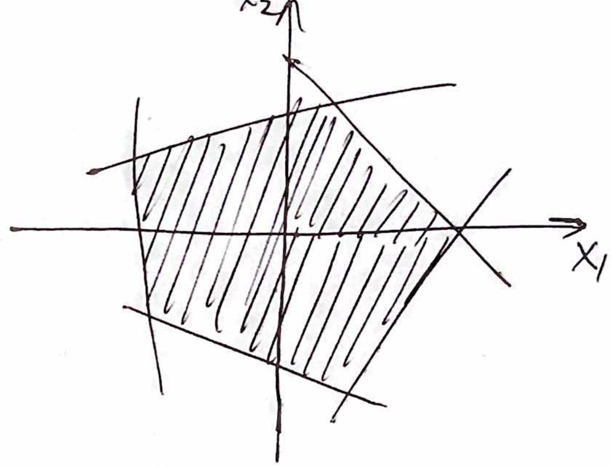
$$= \{x \in \mathbb{R}^n \mid Ax \leq b,\ Cx = d\}$$

where $A = [a_1 \cdots a_m]^T$, $b = [b_1 \cdots b_m]^T$, $C = [c_1 \cdots c_p]^T$, $d = [d_1, \dots, d_p]^T$

- A <u>polytope</u> is a <u>bounded</u> polyhedron



a polyhedron (unbounded) in $\mathbb{R}^2$



a polytope in $\mathbb{R}^2$

- the convex hull of a finite set $\{x_1, \dots, x_k\}$ is a polytope

Norm balls, Ellipsoids, and Norm cones

- A <u>norm ball</u> of radius r and center x_c is defined by the following set for any norm

$$\{x \in \mathbb{R}^n \mid \|x - x_c\| \le r\}$$

  - the <u>Euclidean ball</u> uses the 2-norm

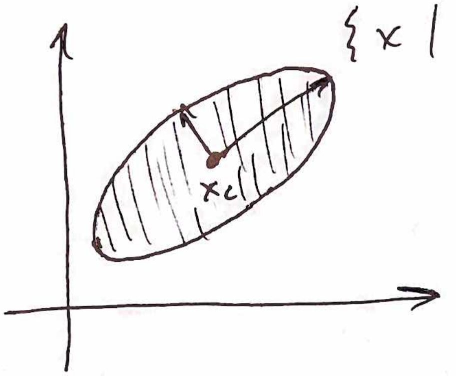
  - An <u>ellipsoid</u> is defined by

$$\{x \in \mathbb{R}^n \mid (x - x_c)^T P^{-1} (x - x_c) \le 1\} \qquad P = P^T \succ 0$$

  i.e. a norm ball of the <u>quadratic norm</u> $\|\cdot\|_{P^{-1}}$

  - norm balls are <u>sublevel sets</u> of norm functions and are all convex

$\{x \mid (x-x_c)^T P^{-1} (x-x_c) \leq 1\}$



- semi axis lenghts are square roots of eigenvalues of P
- Euclidean balls are ellipsoids with $P = r^2 I$

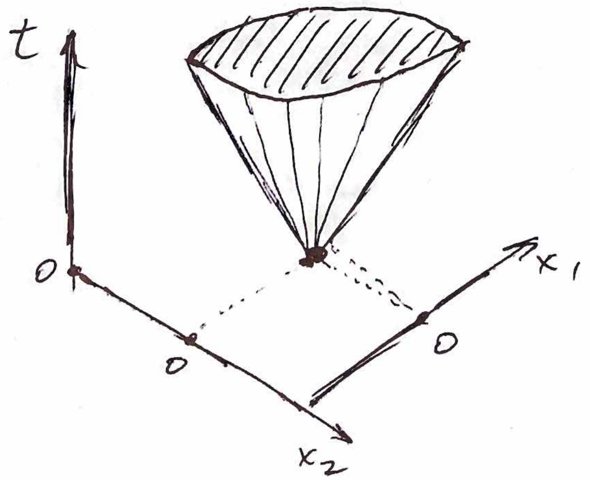
ellipsoid in $\mathbb{R}^2$

- A <u>norm</u> <u>cone</u> is a set of the form

$$\{(x,t) \mid \|x\| \leq t\} \subseteq \mathbb{R}^{n+1}$$

i.e. the <u>epigraph</u> of the norm function

- important special case: the 2-norm cone, aka the <u>second-order</u> <u>cone</u> (or Lorentz cone, or ice-cream cone)



Positive Semidefinite Cone

$$S_+^n = \{ X \in S^n \mid X \succeq 0 \}$$

i.e. the set of all positive semidefinite $n \times n$ matrices

- is a convex cone on S^n

Proof: Let $A, B \succeq 0$ and $\theta \in [0,1]$.
Consider any $y \in \mathbb{R}^n$. Then

$$y^T(\theta A + (1-\theta)B)y = \theta y^T A y + (1-\theta) y^T B y \geq 0$$

Furthermore, for any $A \succeq 0$, $\theta \geq 0$, $y \in \mathbb{R}^n$,

$$y^T(\theta A)y = \theta y^T A y \geq 0 \quad \blacksquare$$

- See BV Fig 2.12

Operations that preserve convexity

- How to prove convexity of a set $C \subseteq \mathbb{R}^n$

  ① apply the definition

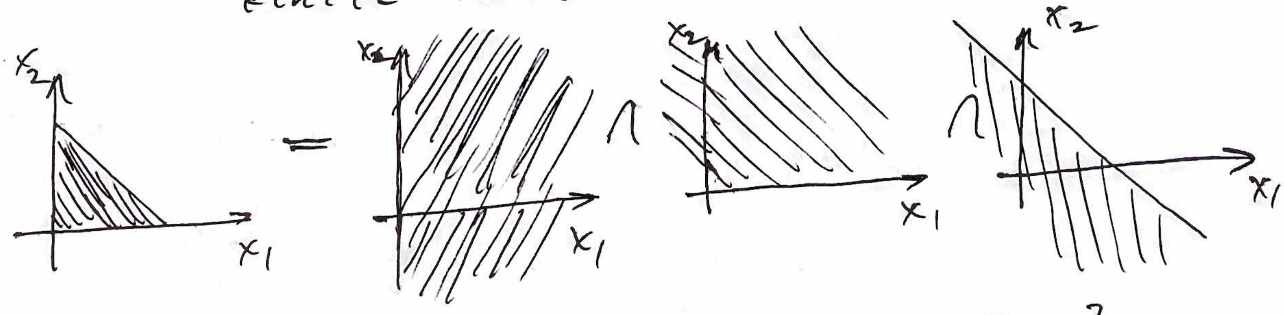
  $$x_1, x_2 \in C,\ \theta \in [0,1] \Rightarrow \theta x_1 + (1-\theta) x_2 \in C$$

  ② show C is obtained from simple convex sets by operations that preserve convexity

- <u>Intersection</u>

  - the intersection of any number of convex sets is convex

    - <u>Ex</u> Polyhedra are intersections of a finite number of half spaces



    - <u>Ex</u> $$S_+^n = \bigcap_{z \neq 0} \{ X \in S^n \mid z^T X z \geq 0 \}$$

      for each $z \neq 0$ $z^T X z$ is <u>linear</u> in X so $\{X \in S^n \mid z^T X z \geq 0\}$ is a halfspace in S^n

- Note: the <u>union</u> of two convex sets is clearly not convex in general

- <u>Affine functions</u>

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

- the <u>image</u> of any convex set under f is convex

$$S \subseteq \mathbb{R}^n \text{ convex} \implies f(S) = \{ f(x) \in \mathbb{R}^m \mid x \in S \} \text{ convex}$$

- the <u>inverse image</u> of a convex set under f is convex

$$C \subseteq \mathbb{R}^m \text{ convex} \implies f^{-1}(C) = \{ x \in \mathbb{R}^n \mid f(x) \in C \} \text{ convex}$$

- examples:
  - scaling, rotation, translation, projection
  - <u>Linear Matrix Inequalities</u>

    $\{ x \in \mathbb{R}^m \mid x_1 A_1 + \cdots + x_m A_m \preceq B \}$
    with $A_i, B \in S^n$ is the <u>inverse image</u> of the positive semidefinite cone under the affine function $f: \mathbb{R}^m \to S^n$ given by
    $$f(x) = B - \left(\sum_i x_i A_i \right)$$

- More generally, images + inverse images of convex sets are preserved under
  - perspective functions $P: \mathbb{R}^{n+1} \to \mathbb{R}^n$

    $$P(x,t) = \frac{x}{t}, \quad \text{dom}(P) = \{(x,t) \mid t > 0\}$$
  - linear-fractional functions $f: \mathbb{R}^n \to \mathbb{R}^m$

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

Generalized Inequalities

A convex cone $K \subseteq \mathbb{R}^n$ is called proper if

- K is closed
- K is solid (has nonempty interior)
- K is pointed (contains no line)

Examples:

- nonnegative orthant: $K = \mathbb{R}^n_+ = \{x \in \mathbb{R}^n \mid x_i \geq 0 \ \forall i\}$
- norm cones: $K = \{(x,t) \in \mathbb{R}^{n+1} \mid \|x\| \leq t\}$
- positive semidefinite cone: $K = S^n_+$
- nonnegative polynomials on $[0,1]$

  $$K = \{x \in \mathbb{R}^n \mid x_1 + x_2 t + x_3 t^2 + \cdots + x_n t^{n-1} \geq 0 \ \forall t \in [0,1]\}$$

- generalized inequality defined by proper cone K

$$x \preceq_K y \iff y - x \in K, \quad x \prec_K y \iff y - x \in \operatorname{int} K$$

Ex

- component wise inequality ($K \in \mathbb{R}^n_+$)

$$x \preceq_{\mathbb{R}^n_+} y \iff x_i \le y_i \quad \forall i$$

- matrix inequality ($K = S^n_+$)

$$X \preceq_{S^n_+} Y \iff Y - X \in S^n_+$$

- many properties of $\preceq_K$ similar to $\le$ on $\mathbb{R}$, but $\preceq_K$ is not in general a linear ordering, only a partial ordering
  - can have $x \npreceq_K y$ and $y \npreceq_K x$
  - $x \in S$ is the minimum element of S wrt $\preceq_K$ if

$$y \in S \implies x \preceq_K y$$

  - $x \in S$ is a minimal element of S wrt $\preceq_K$ if

$$y \in S, \; y \preceq_K x \implies y = x$$

  - see BV Fig 2.17

Separating + Supporting Hyperplane Theorems

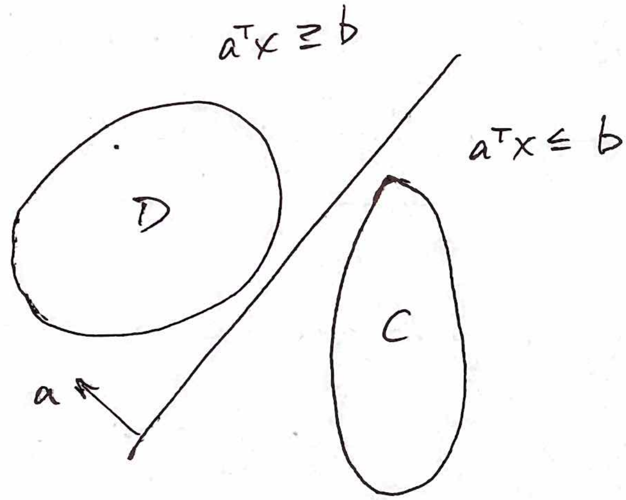
A fundamental result about convex sets:

$C \cap D = \phi$

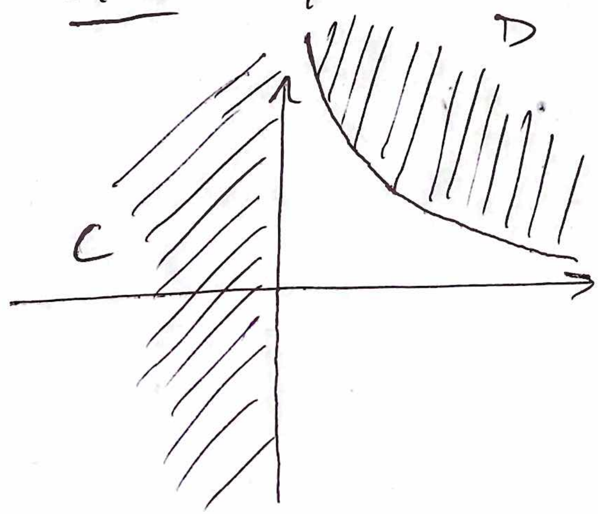
Theorem Let C and D be nonempty disjoint convex sets. Then $\exists a \in \mathbb{R}^n$, $a \neq 0$, $b \in \mathbb{R}$ such that

$$a^T x \leq b \ \forall x \in C \quad \text{and} \quad a^T x \geq b \ \forall x \in D$$

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



Note: Strict separation requires additional assumptions



- Two closed convex sets that can't be strictly separated

A special case w/ strict separation:

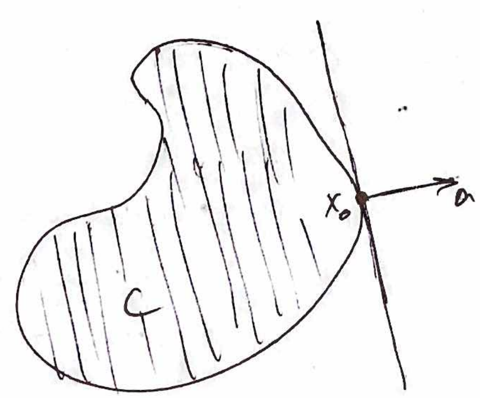
Theorem Let C and D be nonempty closed convex sets in $\mathbb{R}^n$ with at least one of them bounded. Then $\exists a \in \mathbb{R}^n$, $a \neq 0$, $b \in \mathbb{R}$ such that

$$a^T x < b \quad \forall x \in C \quad \text{and} \quad a^T x > b \quad \forall x \in D$$

An important corollary:

Corollary Let $C \subseteq \mathbb{R}^n$ be a closed convex set and $x \in \mathbb{R}^n$ be a point not in C. Then x and C can be strictly separated by a hyperplane.

Definition: A supporting hyperplane to a set C at boundary point x_0 is $\{x \mid a^T x = a^T x_0\}$ where $a \neq 0$ and $a^T x \leq a^T x_0 \ \forall x \in C$



Theorem: If C is convex, then there exists a supporting hyperplane at every boundary point of C.

Dual Cones and Generalized Inequalities

The <u>dual cone</u> of a cone K is the set

$$K^* = \{ y \in \mathbb{R}^n \mid x^T y \geq 0 \quad \forall x \in K \}$$

- examples:
  - $K = \mathbb{R}^n_+ \; : \quad K^* = \mathbb{R}^n_+$
  - $K = S^n_+ \; : \quad K^* = S^n_+$
  - $K = \{ (x,t) \mid \|x\|_2 \leq t \} \; : \quad K^* = \{ (x,t) \mid \|x\|_2 \leq t \}$
  - $K = \{ (x,t) \mid \|x\|_1 \leq t \} \; : \quad K^* = \{ (x,t) \mid \|x\|_\infty \leq t \}$

- Dual cones + generalized inequalities can be used to characterize minimum and minimal elements of sets, and are important in duality theory and describing fundamental tradeoffs (more later)
  - see BV §2.6 for more details