EE445 Mod4-Lec1: Convexity and Convex Sets

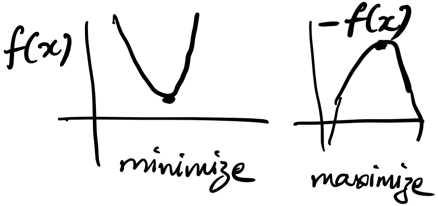
References: [Optimization Models: Calafiore & El Ghaoui] Chapter 8

Topics for Module 4

- Fri May 20 – is in different room : GUG 218 (Guggenheim Hall)

- Lec1: Convexity and Convex Sets
- Lec2: Convex Functions, Smooth Unconstrained Minimization & Gradient Descent
- Lec3: Convex Optimization Problems: ML models I
- Lec4: Convex Optimization Problems: ML models II

Optimization: Overview



A general optimization problem has the form

$$\left[\begin{array}{ll} \text{minimize}_x & f_0(x) \\ \text{subject to} & f_i(x) \le b_i, \quad i = 1, \ldots, \underline{m}, \end{array}\right.$$

e.g., least-norm prob:

$$\left[\begin{array}{ll} \min\limits_{x} & \|x\|^2 \\ \text{s.t.} & Ax = b \end{array}\right.$$

$$\begin{bmatrix} & A & \end{bmatrix} \begin{bmatrix} x \end{bmatrix} = b$$

with components

- $x = \underline{(x_1, \ldots, x_n)}$ - optimization variable $\quad x \in \mathbb{R}^n$
- $f_0 : \mathbf{R}^n \to \mathbf{R}$ - objective function
- $f_i : \mathbf{R}^n \to \mathbf{R}$ - constraint functions; b_i - constraint bounds

$$\underbrace{f_i(x) - b_i}_{\text{can be seen as new constraint function}} \le 0$$

Optimization: Applications

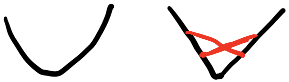
Many applications:

- Data fitting and regression
- Classification (train a classifier)
- Image processing
- Portfolio optimization
- Recommender systems
- Optimal control (design control law for a robot)
- Sensor placement
- Medical treatment planning
- Routing and scheduling
- ...

Optimization: Problem Classes

There are different classes of optimization problems, which can determine a problem's difficulty and solution method:

- Constrained vs. Unconstrained
- Smooth vs. Nonsmooth
- Convex vs. Nonconvex



quote from Prof. Terry Rockafellar (UW Math)

Other important considerations include problem size, special structure such as sparsity, and uncertainty in data or model parameters.

Optimization: Problem Classes

An important class: **convex optimization** problems

"With only a bit of exaggeration, we can say that if you formulate a practical problem as a convex optimization problem, then you have solved the original problem."

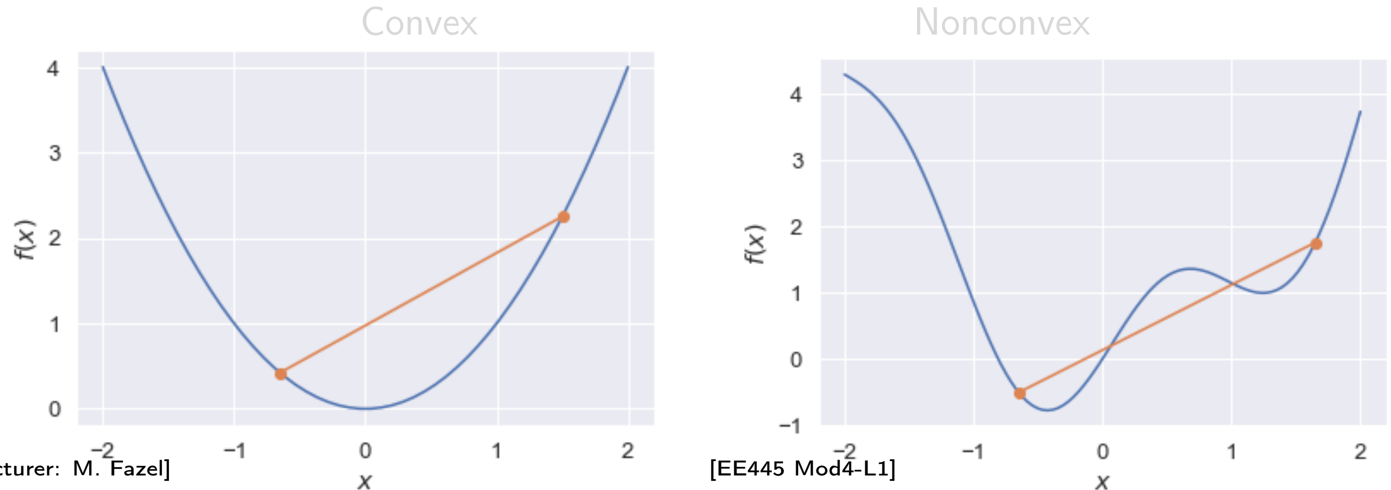
Quote from: S. Boyd and L. Vandenberghe. Convex Optimization. Cambridge University Press, 2004.

Convex Optimization

A convex optimization problem has objective and constraint functions that satisfy the inequality

$$f_i\Big(\lambda x + (1-\lambda)y\Big) \le \lambda f_i(x) + (1-\lambda) f_i(y)$$

for all $x, y \in \mathbf{R}^n$ and all $0 \le \lambda \le 1$.



Convex Optimization

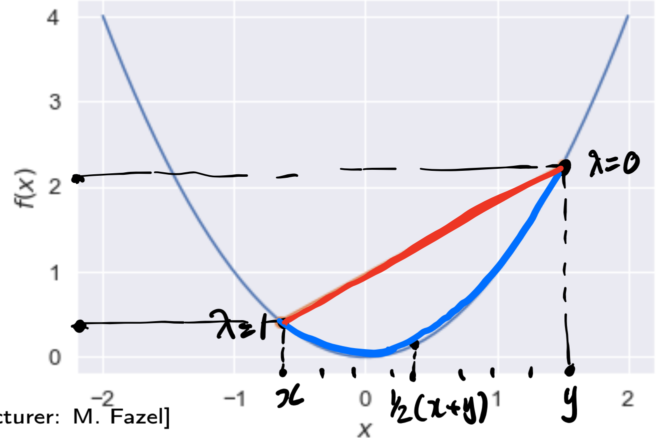
A convex optimization problem has objective and constraint functions that satisfy the inequality

$$f_i\left(\overbrace{\lambda x + (1-\lambda) y}^{\text{mixture of } x,y \,=\, \text{convex combination}}\right) \leq \lambda f_i(x) + (1-\lambda) f_i(y)$$

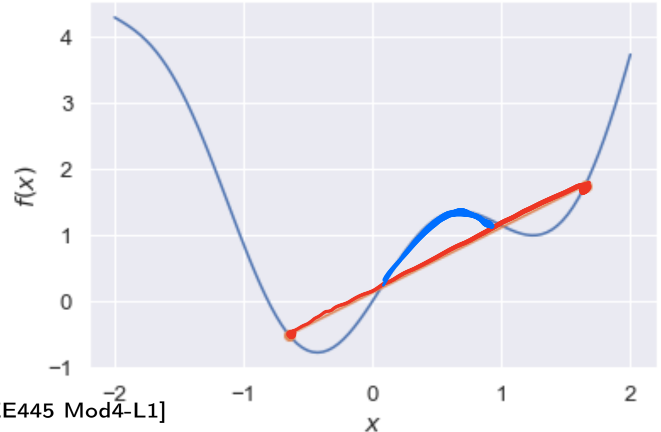
$\lambda = 0 : \quad y$
$\lambda = 1 : \quad x$

for all $x, y \in \mathbf{R}^n$ and all $0 \leq \lambda \leq 1$.

Convex



Nonconvex



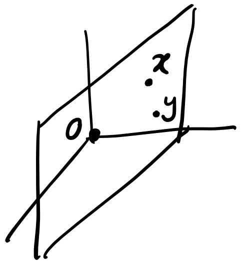
Convex sets

- Convex functions are directly related to convex sets:
  the set defined by
  $$\{x \in \mathbf{R}^n \mid f_i(x) \leq 0, \ i=1,\dots,m\}$$
  is a convex set if the functions f_i are convex (will see more later)
- the def of a convex set is a generalization of the def of a subspace (and affine set), so we start by reviewing these
- we'll then define convex set, ~~convex~~ convex hulls, ...
- examples

Subspaces

linear combination

$S \subseteq \mathbf{R}^n$ is a *subspace* if for $x, y \in S, \quad \lambda, \mu \in \mathbf{R} \implies \lambda x + \mu y \in S$



geometrically: $x, y \in S \Rightarrow$ plane through $0, x, y \subseteq S$

representations

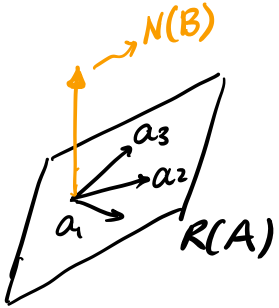
$$\begin{aligned} \text{range}(A) &= \{Aw \mid w \in \mathbf{R}^q\} \\ &= \{w_1 a_1 + \cdots + w_q a_q \mid w_i \in \mathbf{R}\} = \operatorname{span}(a_1, a_2, \ldots, a_q) \end{aligned}$$

where $A = \begin{bmatrix} a_1 & \cdots & a_q \end{bmatrix}$; and

$$\begin{aligned} \text{nullspace}(B) &= \{x \mid Bx = 0\} \\ &= \{x \mid b_1^T x = 0, \ldots, b_p^T x = 0\} \end{aligned}$$

where $B = \begin{bmatrix} b_1^T \\ \vdots \\ b_p^T \end{bmatrix}$

Subspaces



- $S \subseteq \mathbf{R}^n$ is a *subspace* if for $x, y \in S, \quad \lambda, \mu \in \mathbf{R} \implies \lambda x + \mu y \in S$

geometrically: $x, y \in S \Rightarrow$ plane through $0, x, y \subseteq S$

representations

$$\begin{aligned} \text{range}(A) &= \{Aw \mid w \in \mathbf{R}^q\} \\ &= \{w_1 a_1 + \cdots + w_q a_q \mid w_i \in \mathbf{R}\} = \text{span}(a_1, a_2, \ldots, a_q) \end{aligned}$$

where $A = \begin{bmatrix} a_1 & \cdots & a_q \end{bmatrix}$; and

$$\begin{aligned} \text{nullspace}(B) &= \{x \mid Bx = 0\} \\ &= \{x \mid b_1^T x = 0, \ldots, b_p^T x = 0\} \end{aligned}$$

where $B = \begin{bmatrix} b_1^T \\ \vdots \\ b_p^T \end{bmatrix}$

Affine sets

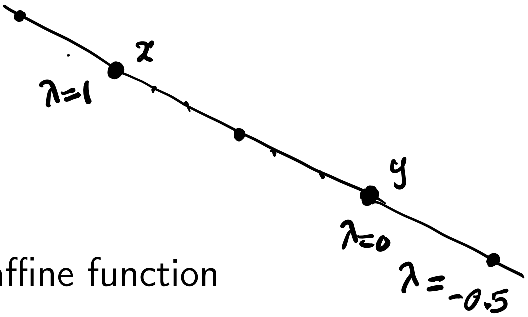
$\mu = 1-\lambda$ $\lambda x + (1-\lambda) y$

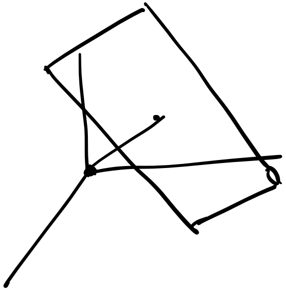
$S \subseteq \mathbf{R}^n$ is *affine* if for $x, y \in S$, $\lambda, \mu \in \mathbf{R}$, $\lambda + \mu = 1 \implies \lambda x + \mu y \in S$

geometrically: $x, y \in S \Rightarrow$ line through $x, y \subseteq S$

$\lambda = -0.5$

$-0.5x + 1.5y$





representations: range of affine function

$$S = \{Az + b \mid z \in \mathbf{R}^q\}$$

via linear equalities

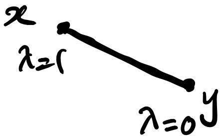
$$\begin{aligned} S &= \{x \mid b_1^T x = d_1, \ldots, b_p^T x = d_p\} \\ &= \{x \mid Bx = d\} \end{aligned}$$

Convex sets

$S \subseteq \mathbf{R}^n$ is a **convex set** if

$\mu = 1-\lambda$ $\qquad$ $\lambda x + (1-\lambda) y$

$$x, y \in S, \quad \lambda, \mu \geq 0, \quad \lambda + \mu = 1 \implies \lambda x + \mu y \in S$$

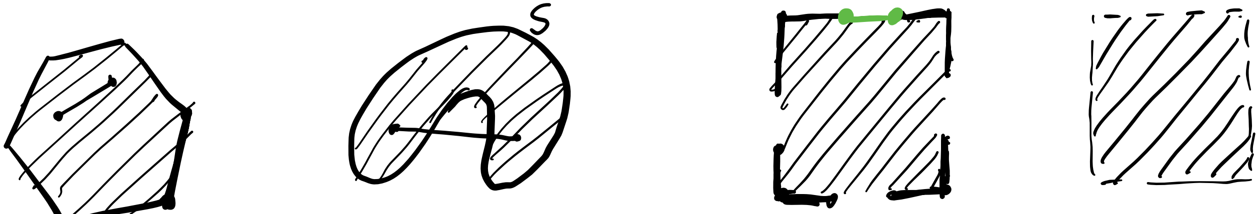


or equivalently, C is convex if

$$x_1, x_2 \in C, \quad 0 \leq \theta \leq 1 \quad \implies \quad \theta x_1 + (1-\theta) x_2 \in C$$

geometrically: $x, y \in S \Rightarrow$ segment $[x, y] \subseteq S$

examples (one convex, two nonconvex sets)



convex $\qquad$ non-convex $\qquad$ nonconvex $\qquad$ convex

Combinations and hulls

consider combinations of more than 2 points:

$y = \theta_1 x_1 + \cdots + \theta_k x_k$ is a

- *linear combination* of $x_1, \ldots, x_k$
- *affine combination* if $\sum_i \theta_i = 1$
- *convex combination* if $\sum_i \theta_i = 1$, $\theta_i \geq 0$
- *conic combination* if $\theta_i \geq 0$

(linear,...) **hull** of S:
set of all (linear, ...) combinations from S

linear hull:	$\mathrm{span}(S)$
affine hull:	$\mathbf{Aff}(S)$
convex hull:	$\mathbf{Co}(S)$
conic hull:	$\mathbf{Cone}(S)$

Convex combination and convex hull

$S = \{x_1, x_2\}$
$Co(S)$

convex combination of $x_1, \ldots, x_k$: any point x of the form

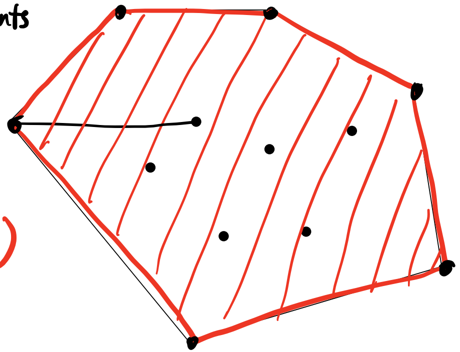
$$\underline{x = \theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_k x_k}$$

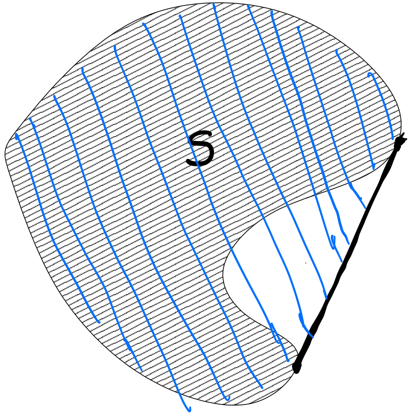
with $\underline{\theta_1 + \cdots + \theta_k = 1}$, $\underline{\theta_i \geq 0}$

convex hull, denoted by $\mathbf{Co}(S)$: set of all convex combinations of points in S

S = set of discrete points

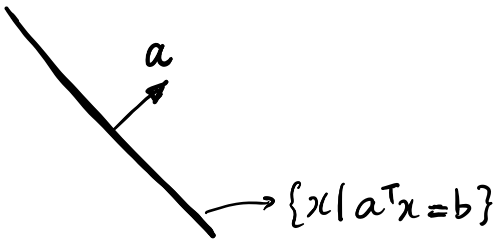
$Co(S)$

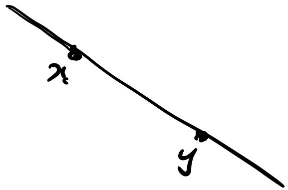




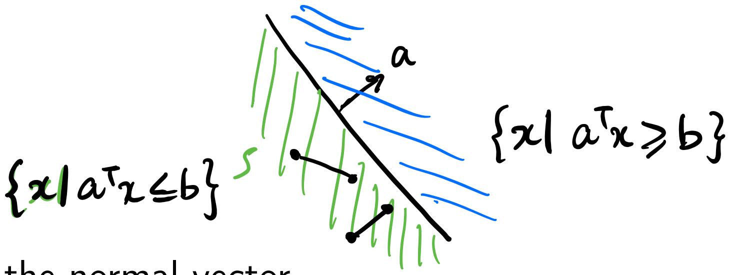
Hyperplanes and halfspaces

hyperplane: set of the form $\{x \mid a^T x = b\}$ $(a \neq 0)$





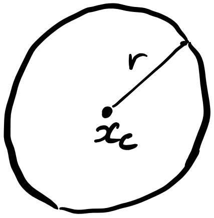
halfspace: set of the form $\{x \mid a^T x \leq b\}$ $(a \neq 0)$



- are halfspaces affine sets?
  No!

- a is the normal vector
- hyperplanes are affine ✓ and convex ✓; halfspaces are convex ✓

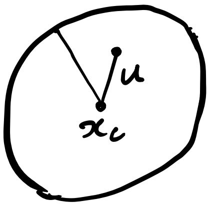
Euclidean balls



(Euclidean) ball with center x_c and radius r:

$$B(x_c, r) = \{x \mid \|x - x_c\|_2 \leq r\} = \{x_c + ru \mid \|u\|_2 \leq 1\}$$

convex set



Ellipsoids

consider special case:

$P = r^2 I \quad r > 0$

$(x-x_c)^T (r^2 I)^{-1} (x-x_c) \le 1$

$\frac{1}{r^2} (x-x_c)^T (x-x_c) \le 1$

$\|x-x_c\|^2 \le r^2$

$\|x-x_c\| \le r$

ellipsoid: set of the form

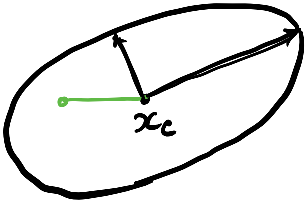
$1\times n \quad n\times n \quad n\times 1$

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

with $P \in \mathbf{S}^n_{++}$ (*i.e.*, P is a symmetric, positive definite matrix)

$S^n_+ : P \succeq 0$

$\rightarrow S^n_{++} : P \succ 0$



Convex set? yes.

Norm balls & norm cones

Recall: norm: a function $\|\cdot\|$ that satisfies $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}_+$

- $\|x\| \geq 0$; $\|x\| = 0$ if and only if $x = 0$
- $\|tx\| = |t| \, \|x\|$ for $t \in \mathbf{R}$
- $\|x + y\| \leq \|x\| + \|y\|$

notation: $\|\cdot\|$ is general (unspecified) norm; $\|\cdot\|_{\mathrm{symb}}$ is a particular norm

norm ball with center x_c and radius r: $\{x \mid \|x - x_c\| \leq r\}$

norm cone: $\{(x,t) \mid \|x\| \leq t\}$

norm balls and cones are convex sets

Norm balls & norm cones

Recall: norm: a function $\|\cdot\|$ that satisfies

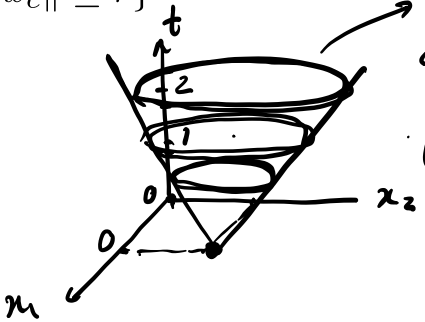
- $\|x\| \geq 0$; $\|x\| = 0$ if and only if $x = 0$
- $\|tx\| = |t|\,\|x\|$ for $t \in \mathbf{R}$
- $\|x + y\| \leq \|x\| + \|y\|$

notation: $\|\cdot\|$ is general (unspecified) norm; $\|\cdot\|_{\text{symb}}$ is a particular norm

norm ball with center x_c and radius r: $\{x \mid \|x - x_c\| \leq r\}$

norm cone: $\{(x,t) \mid \|x\| \leq t\}$

$\{(x,t) \mid \|x\|_2 \leq t\}$ $x \in \mathbb{R}^2$, $t \in \mathbb{R}$



$\{(x,t) \mid \|x\|_2 \leq t\}$

"ice-cream cone"

(x_1, x_2, t)

norm balls and cones are convex sets

Norm balls & norm cones

Recall: norm: a function $\|\cdot\|$ that satisfies

- $\|x\| \geq 0$; $\|x\| = 0$ if and only if $x = 0$
- $\|tx\| = |t| \, \|x\|$ for $t \in \mathbf{R}$
- $\|x + y\| \leq \|x\| + \|y\|$

notation: $\|\cdot\|$ is general (unspecified) norm; $\|\cdot\|_{\text{symb}}$ is a particular norm

norm ball with center x_c and radius r: $\{x \mid \|x - x_c\| \leq r\}$

norm cone: $\{(x, t) \mid \|x\| \leq t\}$

norm balls and cones are convex sets

(can prove this formally using norm properties listed above)

ℓ_p norms

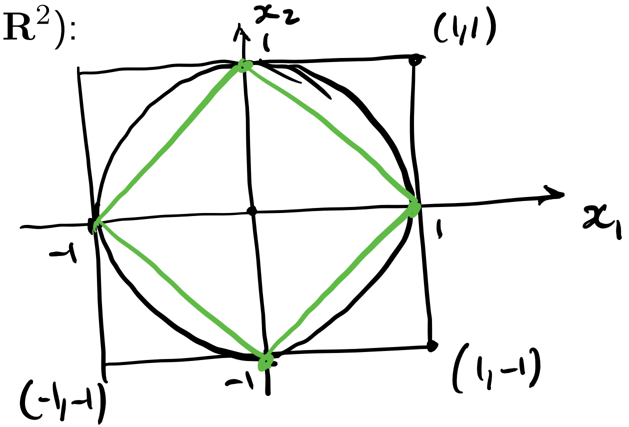
ℓ_p norms on $\mathbf{R}^n$: for $p \geq 1$, $\|x\|_p = \left(\sum_i |x_i|^p\right)^{1/p}$,, for $p = \infty$, $\|x\|_\infty = \max_i |x_i|$

- ℓ_2 norm is Euclidean norm $\|x\|_2 = \sqrt{\sum_i x_i^2}$
- ℓ_1 norm is sum-abs-values $\|x\|_1 = \sum_i |x_i|$
- ℓ_∞ norm is max-abs-value $\|x\|_\infty = \max_i |x_i|$

corresponding norm balls (in $\mathbf{R}^2$):

$\{x \mid \|x\|_\infty \leq 1\}$

$= \{x \mid |x_1| \leq 1, |x_2| \leq 1\}$



$\{x \mid \|x\|_1 \leq 1\}$

$= \{x \mid |x_1| + |x_2| \leq 1\}$

$x_1 + x_2 \leq 1$

$x_1 - x_2 \leq 1$

$-x_1 + x_2 \leq 1$

$-x_1 - x_2 \leq 1$

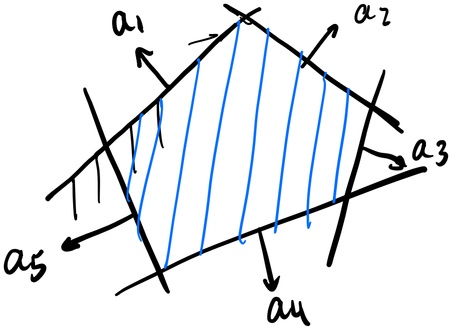
Polyhedra (polytope)

solution set of finitely many linear inequalities and equalities

$$\{x \mid Ax \preceq b, \quad Cx = d\} = \{x \mid a_i^T x \le b_i \ \ i=1,\dots,m, \ \ c_i^T x = d_i \ \ i=1,\dots,p\}$$

($A \in \mathbf{R}^{m\times n}$, $C \in \mathbf{R}^{p\times n}$, $\preceq$ is componentwise inequality)

$i=5$



$a_1^T x \le b_1$

polyhedron is the intersection of a finite number of halfspaces and hyperplanes

Operations that preserve convexity

practical methods for establishing convexity of a set C

1. apply definition

$$\underline{x_1, x_2 \in C}, \quad \underline{0 \leq \theta \leq 1} \quad \Longrightarrow \quad \theta x_1 + (1-\theta) x_2 \in C$$

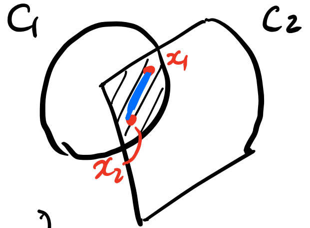
2. show that C is obtained from simple convex sets (hyperplanes, halfspaces, norm balls, ...) by operations that preserve convexity
   - intersection
   - affine functions

Convexity preserved under *intersection*

important property: the intersection of (any number of, *even infinite*) convex sets is convex.

proof: $\forall\ x_1 \in C_1 \cap C_2$

~~examples~~: $\forall\ x_2 \in C_1 \cap C_2$



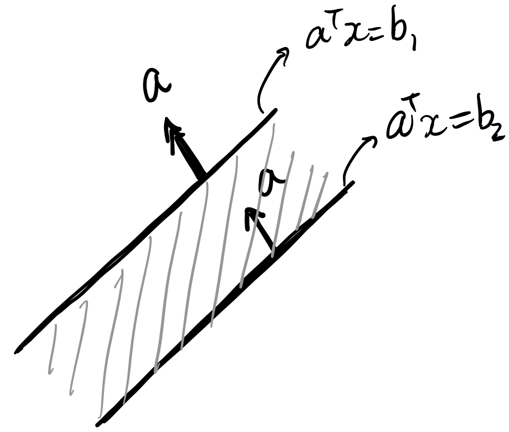
$$\left.\begin{array}{l} x_1, x_2 \in C_1 \ \& \ C_1 \text{ convex set} \Rightarrow \theta x_1 + (1-\theta) x_2 \in C_1, \quad 0 \le \theta \le 1 \\ x_1, x_2 \in C_2 \ \& \ C_2 \quad \text{"} \quad \text{"} \quad \Rightarrow \theta x_1 + (1-\theta) x_2 \in C_2, \quad 0 \le \theta \le 1 \end{array}\right\}$$

$$\Rightarrow \underbrace{\theta x_1 + (1-\theta) x_2}_{\text{line segment } [x_1, x_2]} \in C_1 \cap C_2 \quad 0 \le \theta \le 1$$

example: "Slab"

given $a \in \mathbb{R}^n$, $b_1, b_2 \in \mathbb{R}$,

$\rightarrow \{x \mid a^T x \leq b_1, \; a^T x \leq b_2\}$



convex?

yes, intersection of 2 halfspaces

(using intersection property)

illustration:

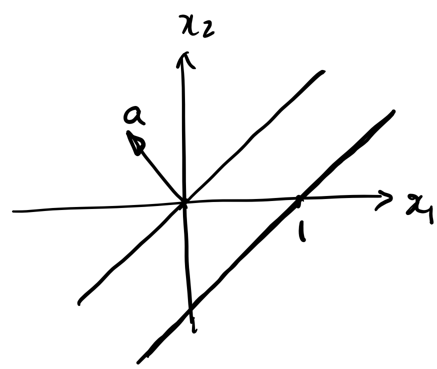
$a = \begin{bmatrix} -1 \\ +1 \end{bmatrix}, \; b_1 = 0$

$a^T x \leq 0$

$\begin{bmatrix} -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \leq 0$

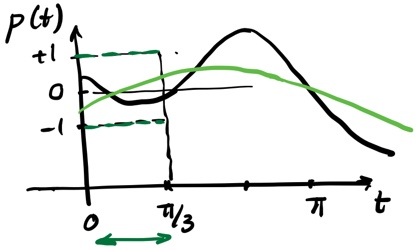
$b_2 = -1 \qquad \begin{bmatrix} -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \leq -1$

$-x_1 + x_2 \leq -1$



Convexity preserved under *intersection*

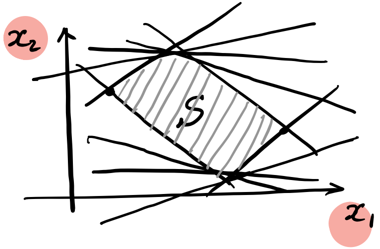
[this ex. taken from: Boyd & Vandenberghe, Convex Optimization, 2004. page 37]



example:

$$S = \{x \in \mathbf{R}^m \mid |p(t)| \leq 1 \text{ for } |t| \leq \pi/3\}$$

where $p(t) = x_1 \cos t + x_2 \cos 2t + \cdots + x_m \cos mt$
for $m = 2$:



$$p(t) = x_1 \cos t + x_2 \cos 2t = \begin{bmatrix} \cos t & \cos 2t \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$S = \{ x \in \mathbb{R}^2 \mid -1 \leq p(t) \leq 1 \quad \text{for all } -\pi/3 \leq t \leq \pi/3 \}$$

$t = t_0$:

$$S_{t_0} = \{ x \in \mathbb{R}^2 \mid -1 \leq \underbrace{(\cos t_0 \quad \cos 2t_0)}_{a_{t_0}} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \leq 1 \} = \{ x \mid -1 \leq a^T x \leq 1 \}$$

change t: get an a_t each time, get a slab. So, S = intersection of (many) slabs, therefore it's convex!

Convexity preserved under *affine function*

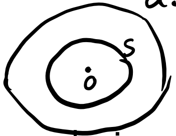
suppose $f : \mathbf{R}^n \to \mathbf{R}^m$ is affine ($f(x) = Ax + b$ with $A \in \mathbf{R}^{m\times n}$, $b \in \mathbf{R}^m$), then the image of a convex set under f is convex

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

$$= \{Ax + b \mid x \in S\} \text{ convex}$$

$\alpha = 2$:

$\{2x \mid x \in S\}$ convex



examples

- scaling, translation
- projection