Administrivia

- HW1 will be out tonight! Have fun ☺

Simplex Algorithm in 1 Slide

Canonical form: $z = z_0 + c_{m+1} x_{m+1} \dots + c_n x_n$

$$
\begin{array}{l}
-z \qquad + c_{m+1} x_{m+1} + \dots + c_j x_j + \dots + c_n x_n = -z_0 \\
x_1 \qquad + a_{1,m+1} x_{m+1} + \dots + a_{1j} x_j + \dots + a_{1n} x_n = b_1 \\
\quad x_2 \qquad \vdots \\
\qquad \ddots \\
\qquad\quad x_m + a_{m,m+1} x_{m+1} + \dots + a_{mj} x_j + \dots + a_{mn} x_n = b_m
\end{array}
$$

If we do pivot in $A_{r,s} > 0$, where $c_s < 0$

$\boxed{X_s \text{ IN} \quad X_r \text{ OUT}}$

New cost value:

$$\tilde{z} = z_0 + \underset{\geq 0}{b_r} \frac{\overset{0 >}{c_s}}{\underset{> 0}{A_{rs}}}$$

CONSTRAINTS

$$
\begin{array}{l}
x_1 = b_1 - A_{1s} x_s \geq 0 \\
\vdots \\
x_m = b_m - A_{ms} x_s \geq 0
\end{array}
$$

$\Downarrow$ $x_s = \text{MIN}\left(\frac{b_i}{A_{is}}\right)$

New b vector:

$$\tilde{b}_j = b_j + b_r \frac{-A_{js}}{A_{rs}} \geq 0 \implies$$

$$\text{LET } r = \underset{\{i \mid A_{is} > 0\}}{\text{ARG MIN}} \frac{b_i}{A_{is}}$$

The full Simplex Algorithm

So far we have assumed that a basic feasible solution in canonical form is available to start the algorithm...

The Simplex Algorithm Phase I

Example

$$\left.\begin{array}{l} 2x_1 + 1x_2 + 2x_3 + x_4 + 4x_5 = z \rightarrow \text{MIN} \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad x \geq 0 \\ \text{s.t } 4x_1 + 2x_2 + 13x_3 + 3x_4 + x_5 = 17 \\ \qquad x_1 + x_2 + 5x_3 + x_4 + x_5 = 7 \end{array}\right\} (*1)$$

Goal: We want to find a feasible solution

Phase I:

(i) Forget the cost function $c^T x$.

(ii) Introduce $X_6, X_7 \geq 0$. [One variable for each row]

(iii) Solve

$$\text{MIN } x_6 + x_7 = w$$

$$x_i \geq 0 \quad i = 1..7$$

$$\text{s.t } \begin{array}{l} 4x_1 + 2x_2 + 13x_3 + 3x_4 + x_5 + x_6 + 0x_7 = 17 \\ x_1 + x_2 + 5x_3 + x_4 + x_5 + 0x_6 + x_7 = 7 \end{array}$$

The Simplex Algorithm Phase I

$$\left.\begin{array}{ll} & \text{MIN } x_6 + x_7 = w \\ & x_i \geq 0 \quad i = 1..7 \\ \text{s.t} & 4x_1 + 2x_2 + 13x_3 + 3x_4 + x_5 + x_6 + 0x_7 = 17 \\ & x_1 + x_2 + 5x_3 + x_4 + x_5 + 0x_6 + x_7 = 7 \end{array}\right\} (*2)$$

Theorem: (*2) has feasible optimal solution such that X_6=X_7=0

iff (*1) has feasible solution

Remarks:

- (*2) is easy to convert to a feasible canonical solution (We will see)
- We can find its optimal solution (X_6=X_7=0) with the Phase II algorithm This is a feasible solution of (*1)

Tableaux

MIN $X_6 + X_7 = w$

$x_i \geq 0 \quad i = 1..7$

s.T $4x_1 + 2x_2 + 13x_3 + 3x_4 + x_5 + x_6 + 0x_7 = 17$

$x_1 + x_2 + 5x_3 + x_4 + x_5 + 0x_6 + x_7 = 7$

(×2)

Basic variable	Objective (-w)	X_1	X_2	X_3	X_4	X_5	X_6	X_7	RHS
	1	0	0	0	0	0	1	1	0
	0	4	2	13	3	1	1	0	17
	0	1	1	5	1	1	0	1	7

It's easy to convert this to a feasible canonical form:

Basic variable	Objective (-w)	X_1	X_2	X_3	X_4	X_5	X_6	X_7	RHS
-w	1	-5	-3	-18	-4	-2	0	0	-24
X_6	0	4	2	13	3	1	1	0	17
X_7	0	1	1	5	1	1	0	1	7

CANONICAL FORM

Tableaux

Now, we can run the Phase 2 algorithm on this table to get a feasible solution of (*1):

B. var	Obj. (-w)	X_1	X_2	X_3	X_4	X_5	X_6	X_7	RHS
-w	1	-5	-3	-18	-4	-2	0	0	-24
X_6	0	4	2	13	3	1	1	0	17
X_7	0	1	1	5	1	1	0	1	7

-18 IS THE SMALLEST $C_j < 0$ $\Rightarrow$ PIVOT IN COLUMN 3

$\frac{17}{13} < \frac{7}{5} \Rightarrow \begin{cases} \text{PIVOT IN } 13 \\ X_3 \text{ IN} \\ X_6 \text{ OUT} \end{cases}$

Tableaux

Now, we can run the Phase 2 algorithm on this table to get a feasible solution of (*1):

B. var	Obj. (-w)	X_1	X_2	X_3	X_4	X_5	X_6	X_7	RHS
-w	1	-5	-3	-18	-4	-2	0	0	-24
X_6	0	4	2	13	3	1	1	0	17
X_7	0	1	1	5	1	1	0	1	7

$\Downarrow$

B. var	Obj (-w)	X_1	X_2	X_3	X_4	X_5	X_6	X_7	RHS
-w	1	-5+18/13*4	-3+18/13*2	0	-4+18/13*3	-2+18/13	18/13	0	-24+18/13*7
X_3	0	4/13	2/13	1	3/13	1/13	1/13	0	17/13
X_7	0	1-5/13*4	1-5/13*2	0	1-5/13*3	1-5/13	-5/13	1	7-5/13*7

Tableaux

B. var	Obj (-w)	X_1	X_2	X_3	X_4	X_5	X_6	X_7	RHS
-w	1	-5+18/13*4	-3+18/13*2	0	-4+18/13*3	-2+18/13	18/13	0	-24+18/13*7
X_3	0	4/13	2/13	1	3/13	1/13	1/13	0	17/13
X_7	0	1-5/13*4	1-5/13*2	0	1-5/13*3	1-5/13	-5/13	1	7-5/13*7

Let us simplify this Table a little:

B. var	Obj (-w)	X_1	X_2	X_3	X_4	X_5	X_6	X_7	RHS
-w	1	-7/13	-3/13	0	2/13	-8/13	18/13	0	-6/13
X_3	0	4/13	2/13	1	3/13	1/13	1/13	0	17/13
X_7	0	1-7/13	3/13	0	-2/13	8/13	-5/13	1	6/13

PIVOT

-8/13 IS THE SMALLEST AMONG ALL $c_j < 0 \Rightarrow x_5$ IN

$$\frac{6/13}{8/13} < \frac{17/13}{1/13} \Rightarrow X_7 \text{ OUT}$$

Tableaux

B. var	Obj (-w)	X_1	X_2	X_3	X_4	X_5	X_6	X_7	RHS
-w	1	0	0	0	0	0	1	1	0
X_3	0	3/8	1/8	1	1/4	0	-1/8	0	5/4
X_5	0	-7/8	3/8	0	-1/4	1	-5/8	1/8	3/4

All the relative costs are nonnegative) optimal feasible solution.

Phase I is finished.

$$\left.\begin{array}{l} -w + x_6 + x_7 = 0 \\ x_3 = 5/4 \\ x_5 = 3/4 \\ x_1, x_2, x_4, x_6, x_7 = 0 \end{array}\right\}$$

FEASIBLE BASIC SOLUTION OF THE ORIGINAL (*1) PROBLEM

$\Downarrow$

PHASE II CAN BE STARTED

Tableaux

B. var	Obj (-w)	X_1	X_2	X_3	X_4	X_5	X_6	X_7	RHS
-w	1	0	0	0	0	0	1	1	0
X_3	0	3/8	1/8	1	1/4	0	-1/8	0	5/4
X_5	0	-7/8	3/8	0	-1/4	1	-5/8	1/8	3/4

ORIGINAL COST FUNCTION: $z = 2x_1 + 1x_2 + 2x_3 + x_4 + 4x_5 \rightarrow \min_{x \geq 0}$

B. var	Obj (-z)	X_1	X_2	X_3	X_4	X_5	X_6	X_7	RHS
-z	1	2	1	2	1	4	0	0	0
X_3	0	3/8	1/8	1	1/4	0	-1/8	0	5/4
X_5	0	-7/8	3/8	0	-1/4	1	-5/8	1/8	3/4

Tableaux

B. var	Obj (-z)	X_1	X_2	X_3	X_4	X_5	X_6	X_7	RHS
-z	1	2	1	2	1	4	0	0	0
X_3	0	3/8	1/8	1	1/4	0	-1/8	0	5/4
X_5	0	-7/8	3/8	0	-1/4	1	-5/8	1/8	3/4

-2x

-4x

Make it to canonical form and continue with Phase II...

Do not make pivots in the column of X_6 and X_7

Simplex Algorithm with Matlab

```
f = [-5  -4  -6]';

A =  [ 1 -1  1
       3  2  4
       3  2  0 ];

b = [20  42 30]';

lb = zeros(3,1);

options = optimset('LargeScale','off','Simplex','on');

[x,fval,exitflag,output,lambda] = linprog(f,A,b,[],[],lb,[],[],options);
```

Relevant Books

- Luenberger, David G. *Linear and Nonlinear Programming*. 2nd ed. Reading, MA: Addison Wesley, 1984. ISBN: 0201157942.

- Bertsimas, Dimitris, and John Tsitsiklis. *Introduction to Linear Optimization*. Athena Scientific Press, 1997. ISBN: 1886529191.

- Dantzig, Thapa: Linear Programming

Summary

- Linear programs:
  - standard form,
  - canonical form
- Solutions:
  - Basic, Feasible, Optimal, Degenerate
- Simplex algorithm:
  - Phase I
  - Phase II
- Applications:
  - Pattern classification

Convex Optimization
CMU-10725

4. Convexity

Barnabás Póczos & Ryan Tibshirani

Goal of this lecture

- **Review of Convex sets & Convex functions**
  - Definition
  - Examples
  - Basic properties

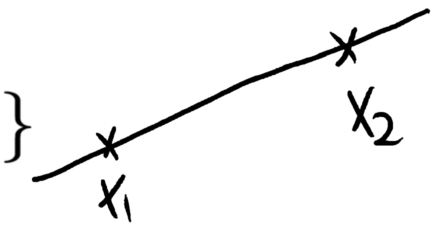
Books to Read:

- **Boyd and Vandenberghe**: Convex Optimization, Chapters 2 & 3
- **Rockafellar**: Convex Analysis

Line and Line Segments

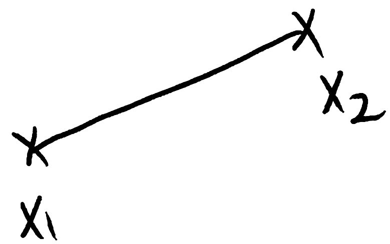
Definition [Line]:

$$\{x \in \mathbb{R}^n | x = \theta x_1 + (1-\theta) x_2, \theta \in \mathbb{R}\}$$



Definition [Line segment]: LINE + $\theta \in [0,1]$!

$$\{x \in \mathbb{R}^n | x = \theta x_1 + (1-\theta) x_2, \theta \in [0,1]\}$$



Affine Sets

Definition [Affine set]:

A set C is affine if for any $x_1, x_2 \in C$
the line through x_1 and x_2 is in C,
i.e. $\underline{\theta x_1 + (1-\theta) x_2 \in C}$, $(\theta \in \mathbb{R})$

Definition [Affine hull of set C]:

$$\mathrm{AFF}[C] = \left\{ \theta_1 x_1 + \ldots + \theta_l x_l \;\middle|\; \begin{array}{l} x_i \in C \\ \sum_{i=1}^{l} \theta_i = 1 \end{array} \right\}$$

Theorem [Affine hull]:

The Aff[C] is the smallest affine set that contains C

Affine Sets Example

Example [Solutions of linear equations]:

The solution set of a system of linear equations is an affine set

Solution set:

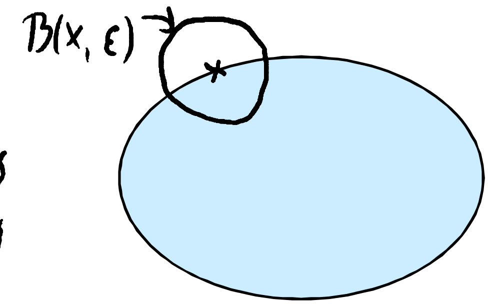
$$C = \{x \mid Ax = b\} \quad A \in \mathbb{R}^{m \times n},\ b \in \mathbb{R}^{m}$$

Proof:

IF

$$\left.\begin{array}{l} Ax_1 = b \\ Ax_2 = b \\ \theta \in \mathbb{R} \end{array}\right\} \Rightarrow A[\theta x_1 + (1-\theta) x_2] = b$$

Boundaries



Definition [x on boundary of C (∂C)]:

FOR SMALL ENOUGH $\varepsilon > 0$: $B(x, \varepsilon) \cap C \neq \emptyset$

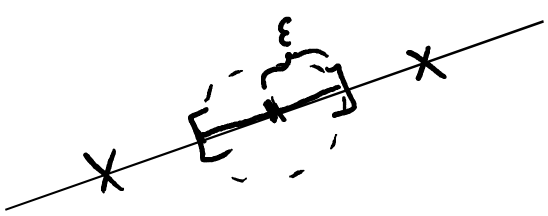
& $B(x, \varepsilon) \cap C^c \neq \emptyset$

Definition [x in interior of C]:

$B(x, \varepsilon) \subset C$ FOR SMALL ENOUGH ε

Definition [*relative* interior (rel int C)]:

REL INT $C = \{ x \in C \mid B(x, \varepsilon) \cap \text{AFF } C \subseteq C$ FOR SOME $\varepsilon > 0 \}$



Boundaries

Definition [closure of C (cl C)]:

$$\mathrm{cl}\, C = C \cup \partial C$$

Definition [*relative* boundary of C (rel ∂C)]:

$$\mathrm{cl}\, C \setminus \mathrm{REL\ INT}\, C$$

Open and Closed Sets

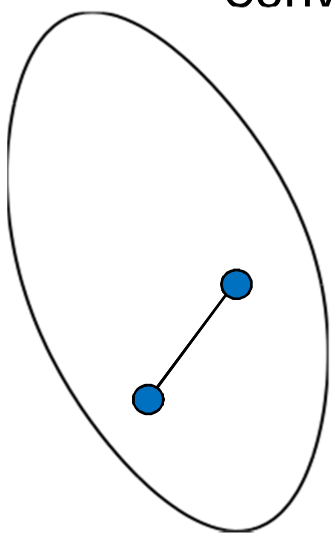
Definition [C closed]: $\partial C \subset C$

Definition [C open]: $\partial C \cap C = \emptyset$

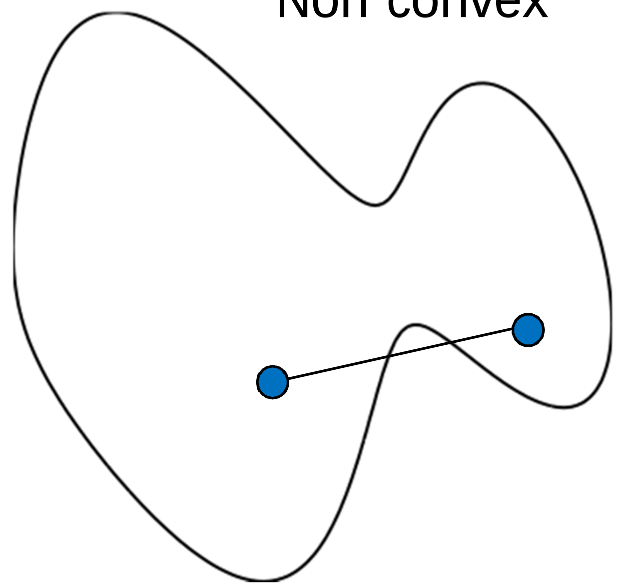
Definition [C compact]: CLOSED AND BOUNDED [IN $\mathbb{R}^n$]

Convex sets

Convex



Non convex



Definition [Convex set]:

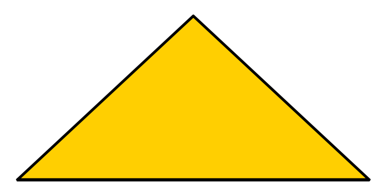
SET $C \subset \mathbb{R}^d$ IS CONVEX $\Longleftrightarrow$ $\forall x_1, x_2 \in C$ $\forall \theta \in [0,1]$

$\theta x_1 + (1-\theta) x_2 \in C$

Definition [Strictly convex set]:

$\theta x_1 + (1-\theta) x_2 \in \mathrm{INT}(C)$ $\forall x_1 \neq x_2 \in C$ $\theta \in (0,1)$

Example [Convex, but not strictly convex]:



Examples of Convex Sets

- empty set: $\emptyset$
- singleton set: $\{x_0\}$
- complete space: $\mathbb{R}^d$

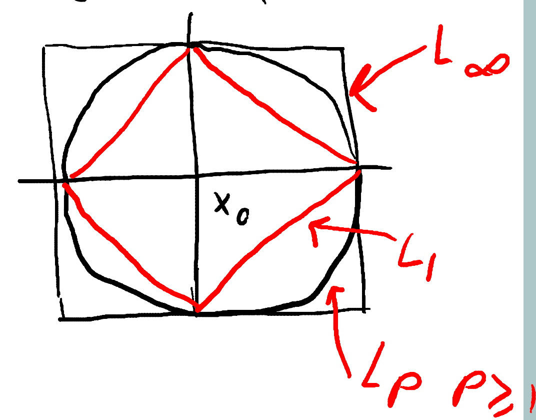
- lines: $\{x \mid x = \theta x_1 + (1-\theta) x_2,\ \theta \in \mathbb{R}\}$
- line segments: $\{x \mid x = \theta x_1 + (1-\theta) x_2,\ \theta \in [0,1]\}$
- hyperplanes: $\{x \in \mathbb{R}^n \mid a^T x = b\}$ $a \in \mathbb{R}^n$, $b \in \mathbb{R}$
- halfspaces: $\{x \in \mathbb{R}^n \mid a^T x \le b\}$ $a \in \mathbb{R}^n$, $b \in \mathbb{R}$

Examples of Convex Sets

- Euclidian balls:

$$\{x \in \mathbb{R}^n \mid \|x - x_0\|_2 \leq r\} = \mathcal{B}(x_0, r)$$

CONVEX



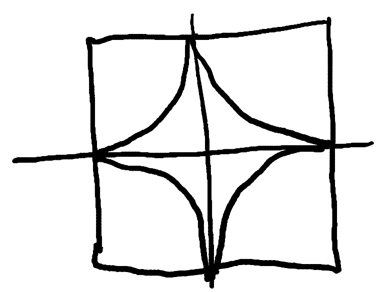
- L_p balls, $p \geq 1$

$$\{x \in \mathbb{R}^n \mid \|x - x_0\|_p \leq r\}$$

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}$$

$$\|x\|_\infty = \max_{i=1..n} |x_i|$$

- L_p balls $0<p<1$



NOT CONVEX!

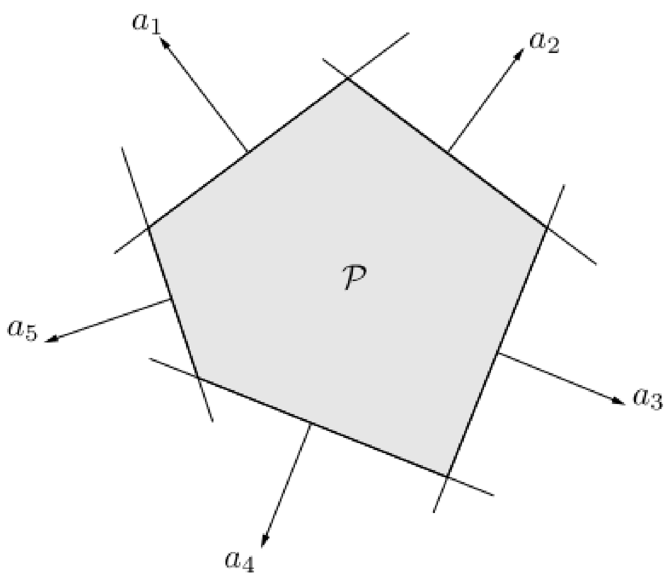
Examples of Convex Sets

- Polyhedron: the solution set of a finite number of linear equalities and inequalities

$$\mathcal{P} = \{ x \mid a_j^T x \leq b_j, \ j = 1, \ldots, m, \quad c_j^T x = d_j, \ j = 1, \ldots, p \}$$

Matrix notation: $\mathcal{P} = \{ x \mid Ax \leq b, \ Cx = d \}$

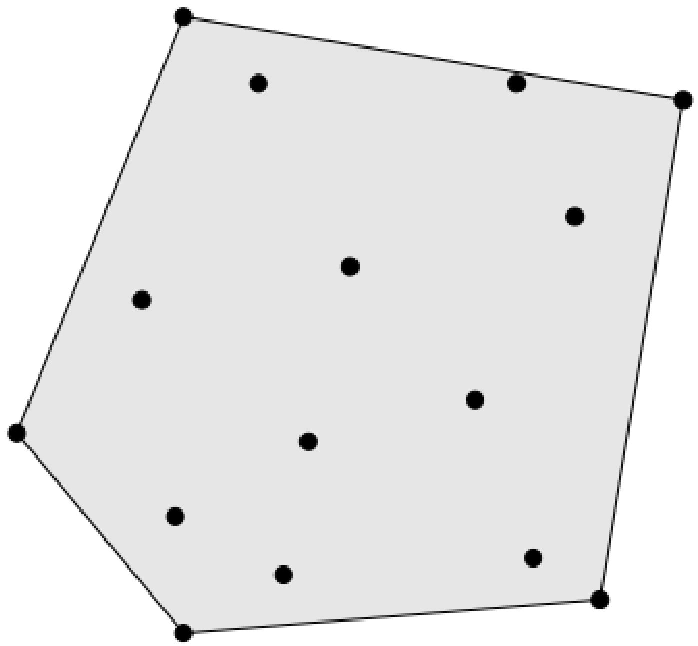
- Polytope: bounded polyhedron

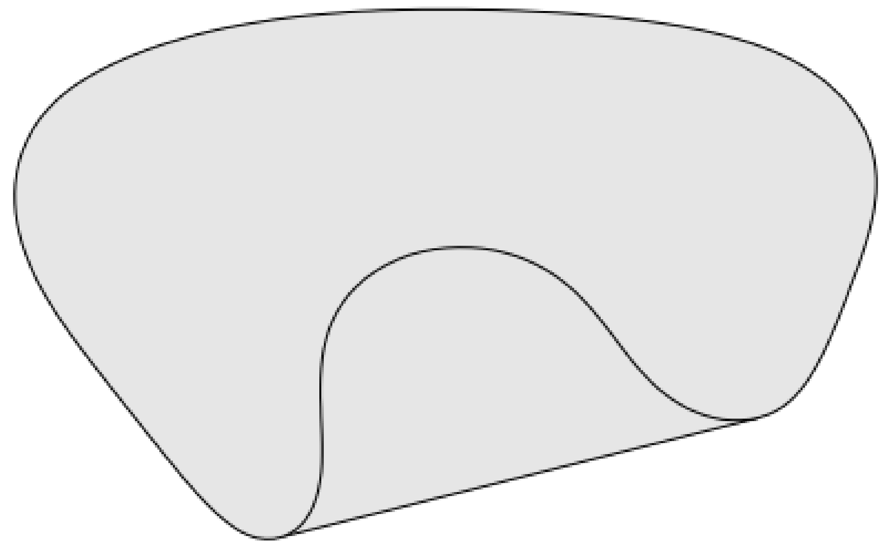


Intersection of halfspaces

& hyperplanes

Convex hull





Definition [Convex hull]:

$$\mathrm{conv}[C] = \left\{ \theta_1 x_1 + \ldots + \theta_k x_k \;\middle|\; x_i \in C,\ \theta_i \geq 0\ \forall i = 1 \ldots k,\ \sum_{i=1}^{k} \theta_i = 1,\ k \in \mathbb{Z}_+ \right\}$$

Convex hull

Convex hull properties:

CONV[C] IS THE SMALLEST CONVEX SET THAT CONTAINS C

CONV[C] IS CONVEX

$C \subseteq \text{conv}(C)$

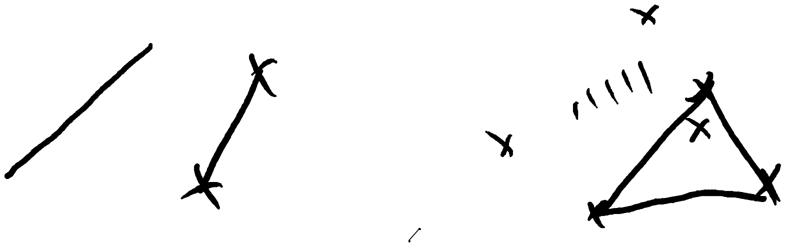
$\forall C, C'$ CONVEX SETS. IF $C \subset C' \Rightarrow \text{conv}[C] \subseteq C'$

Examples of Convex Sets

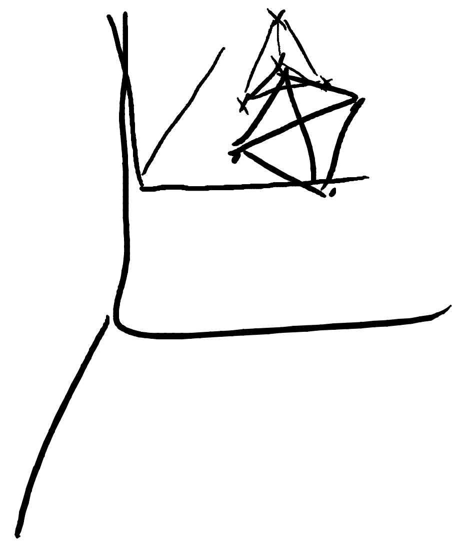
- Simplex:

$$\mathrm{conv}\{\underbrace{v_0, \ldots, v_k}_{k+1}\} = \left\{ \sum_{i=0}^{k} \theta_i v_i \quad \theta_i \geq 0 \quad \sum_{i=0}^{k} \theta_i = 1 \right\}$$

$v_1 - v_0, \ldots, v_k$ 2D



3D



Convex Combination

$$\text{CONV}[C] = \{\theta_1 x_1 + \ldots + \theta_k x_k \mid x_i \in C,\ \theta_i \geq 0\ \forall i = 1 \ldots k,\ \sum_{i=1}^{k} \theta_i = 1,\ k \in \mathbb{Z}_+\}$$

Infinite many sums

C CONVEX $\Rightarrow$ $\underbrace{\sum_{i=1}^{\infty} \theta_i x_i}_{\text{ASSUMING THIS SERIES CONVERGES}} \in C$

$x_1, x_2, \ldots \in C$

$\theta_1, \theta_2, \ldots \geq 0$

$\sum_{i=1}^{\infty} \theta_i = 1$

Integrals (Expected value)

$P : \mathbb{R}^n \to \mathbb{R}$ DENSITY FUNCTION $\Rightarrow$ $\underbrace{\int_C p(x)\, x\, dx}_{\text{ASSUMING THIS INTEGRAL EXIST}} \in C$

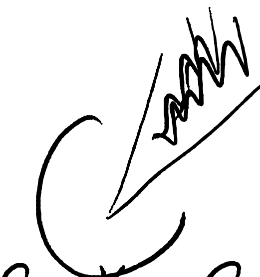
C CONVEX SET

$\mathbb{E}_p[X] \in C$

Cones

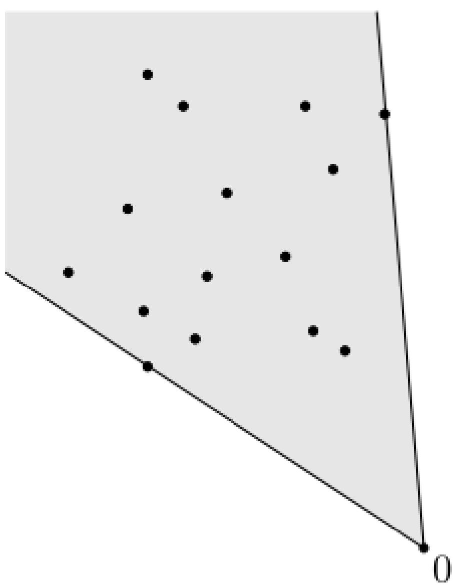
Definition [Cone]:

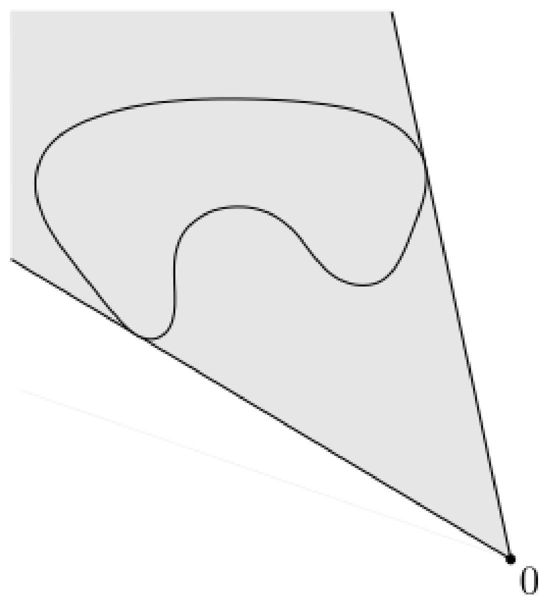
C IS CONE $\Longleftrightarrow \left\{ \begin{array}{l} x \in C \\ \theta \geq 0 \end{array} \Rightarrow \theta x \in C \right\}$



Definition [Convex Cone]: CONE & CONVEX

Definition [Conic hull]: $\text{CONE}[C] = \{ x \mid x = \theta_1 x_1 + \ldots + \theta_k x_k,\ \theta_i \geq 0,\ x_i \in C \}$





Example: PSD matrices

Definition [Positive semi definite matrix]:

$A \in \mathbb{R}^{n \times n}$ MATRIX IS PSD $\Longleftrightarrow \begin{cases} x^T A x \geq 0 \quad \forall x \in \mathbb{R}^n \\ A = A^T \end{cases}$

Theorem [eigenvalues]:

A symmetric matrix A is positive definite iff all its eigenvalues are positive

Partial ordering of square matrices**:**

For arbitrary square matrices M, N we write $M \geq N$

if $M - N \geq 0$; i.e., $M - N$ is positive semi-definite.

Example: PSD matrices

Theorem [Cone of PSD matrices]:

The set of symmetric, PSD matrices form a convex cone:

$$S_+^n = \{ A \in \mathbb{R}^{n \times n} \mid A \succeq 0 \}$$

Proof:

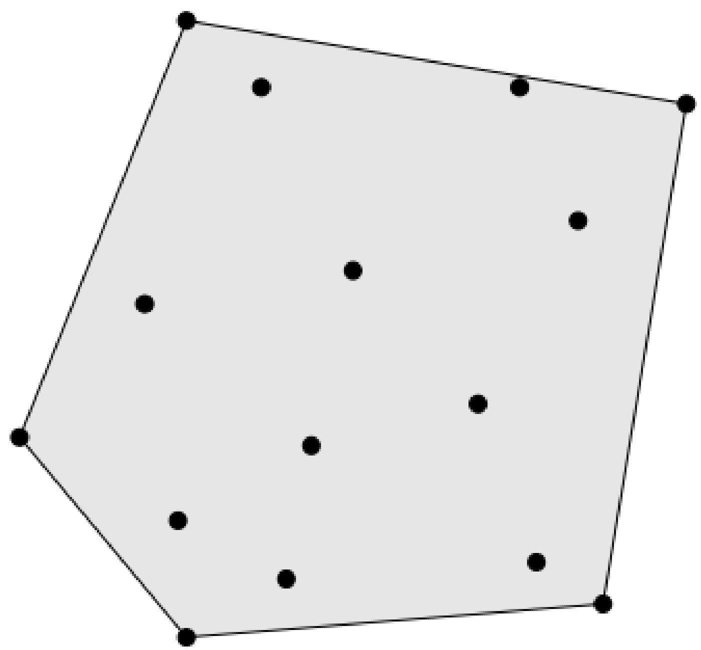
$$A, B \in S_+^n$$

$$\Rightarrow \quad x^T [\theta A + (1-\theta) B] x = \theta \underbrace{x^T A x}_{\geq 0} + (1-\theta) \underbrace{x^T B x}_{\geq 0} \geq 0$$

Convex set representation with convex hull

Theorem: [Representation of a closed convex set with a convex hull]

LET $S \subseteq \mathbb{R}^n \Rightarrow C = cl(\text{conv } S)$ IS THE CONVEX HULL OF POSSIBLY INFINITE MANY POINTS IN S

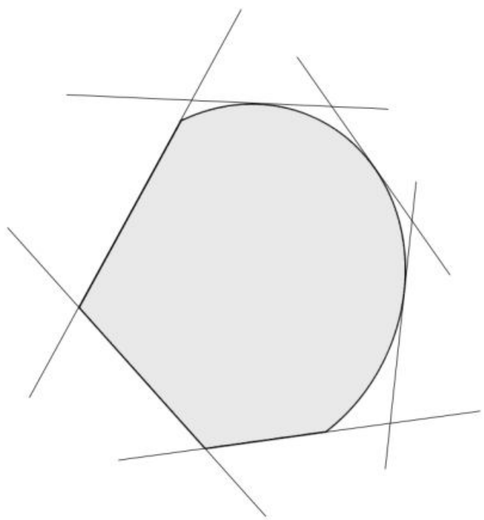


Dual representation

Theorem: [Representation of a closed convex set with half spaces]

LET $S \subseteq \mathbb{R}^n$. $\Rightarrow$ $\overbrace{\mathrm{cl}(\mathrm{conv}\, S)}^{C}$ IS THE INTERSECTION OF ALL THE CLOSED HALF-SPACES WHICH CONTAINS S

$\{x \mid a_i^T x + b \leq 0\}$



Convexity-preserving set operations

- Translation $C + b$
- Scaling αC
- Intersection C, D CONVEX $\Rightarrow C \cap D$ CONVEX

CAN BE EXTENDED TO INFINITE NUMBER OF SETS

IF S_α IS A CONVEX SET $\forall \alpha \in \mathcal{A} \Rightarrow \bigcap_{\alpha \in \mathcal{A}} S_\alpha$ IS CONVAX

Convexity-preserving set operations

- Affine function
  - E.g. projection, dropping coordinates

If $C \subset \mathbb{R}^n$ convex, $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^n$

$\Rightarrow AC + b = \{Ax + b \mid x \in C\} \subseteq \mathbb{R}^m$ is convex.

- Set sum

$$C_1 + C_2 = \{c_1 + c_2 \mid c_1 \in C_1,\ c_2 \in C_2\}$$

- Direct sum

$$C_1 \times C_2 = \{(c_1, c_2) \in \mathbb{R}^{n+m},\ c_1 \in C_1,\ c_2 \in C_2\}$$

Convexity-preserving set operations

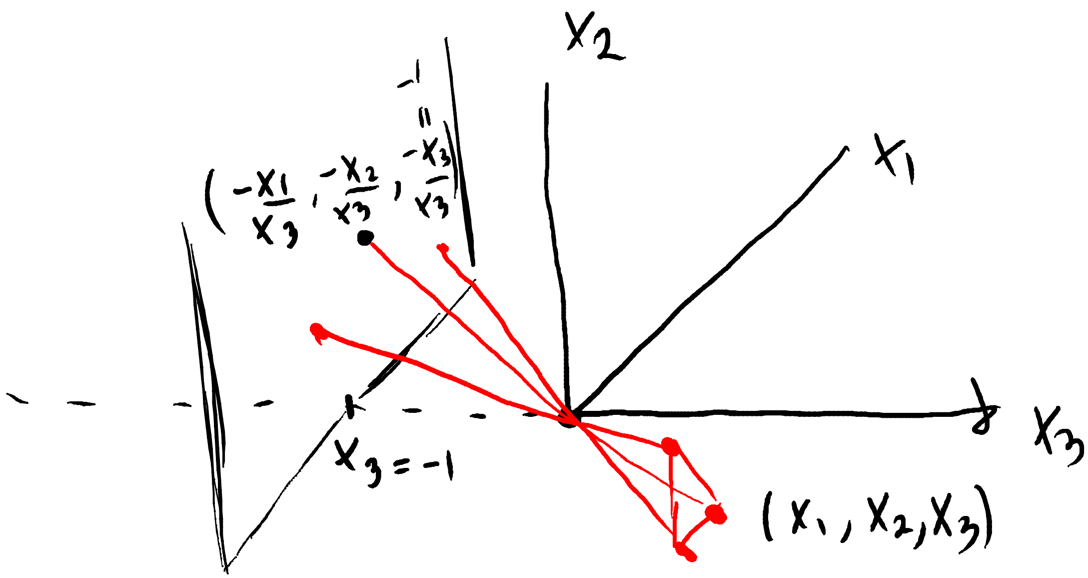
- Perspective projection (pinhole camera)

IF $C \subset \mathbb{R}^n \times \mathbb{R}_{++}$ IS A CONVEX SET, THEN

$P(C)$ IS ALSO CONVEX

$$P(x) = P(\underbrace{x_1, x_2, \ldots, x_n}_{z}, t) = (x_1/t, x_2/t, \ldots, x_n/t) \in \mathbb{R}^n$$

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



Convexity-preserving set operations

- Linear-fractional function

  (perspective function with affine function)

$$f(x) = \frac{Ax+b}{c^Tx+d} \qquad A \in \mathbb{R}^{m\times n} \quad c \in \mathbb{R}^n$$
$$b \in \mathbb{R}^m \quad d \in \mathbb{R}$$

$$\text{Dom } f = \{x \mid c^Tx + d > 0\}$$

Theorem: [Image of Linear fractional function]

$$C \subset \mathbb{R}^n \text{ is convex} \Rightarrow f(C) \subset \mathbb{R}^m \text{ is convex}$$