Administrivia

Simplex Algorithm in 1 Slide

Canonical form:

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

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:

$$\tilde{z} = z_0 + \underbrace{b_r}_{\geq 0} \frac{\overbrace{c_s}^{< 0}}{\underbrace{A_{rs}}_{> 0}}$$

CONSTRAINTS

$$x_1 = b_1 - A_{1s}x_s \geq 0$$

$$\vdots$$

$$x_m = b_m - A_{ms}x_s \geq 0$$

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

New b vector:

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

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

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

$$2x_1 + 1x_2 + 2x_3 + x_4 + 4x_5 = z \rightarrow \text{MIN}, \quad x \geq 0$$

$$\text{s.t.} \quad 4x_1 + 2x_2 + 13x_3 + 3x_4 + x_5 = 17$$

$$x_1 + x_2 + 5x_3 + x_4 + x_5 = 7$$

$(*1)$

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

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

$$\text{s.t.} \quad 4x_1 + 2x_2 + 13x_3 + 3x_4 + x_5 + x_6 + 0x_7 = 17$$

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

The Simplex Algorithm Phase I

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

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 $\quad 4x_1 + 2x_2 + 13x_3 + 3x_4 + x_5 + x_6 + 0x_7 = 17$

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

$\}\ (\times 2)$

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

$\frac{6/13}{8/13} < \frac{17/13}{1/13} \Rightarrow x_7$ 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 $\Rightarrow$ optimal feasible solution.

Phase I is finished.

$$-w + x_6 + x_7 = 0$$
$$x_3 = 5/4$$
$$x_5 = 3/4$$
$$x_1, x_2, x_4, x_6, x_7 = 0$$

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: $z = 2x_1 + 1x_2 + 2x_3 + x_4 + 4x_5 \rightarrow \underset{x \geq 0}{\min}$

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

-2×

-4×

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 Part I

Barnabás Póczos & Ryan Tibshirani

MACHINE LEARNING DEPARTMENT
Carnegie Mellon.
School of Computer Science

Goal of this lecture

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

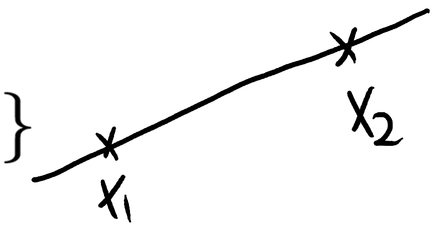
Books to Read:

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

Line and Line Segments

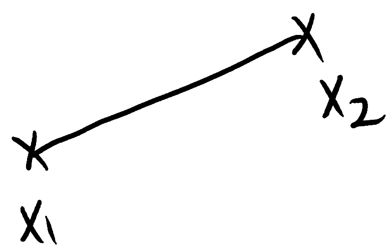
Definition [Line]:

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



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. $\theta x_1 + (1-\theta)x_2 \in C$, $(\theta \in \mathbb{R})$

Definition [Affine hull of set C]:

$C \subseteq \mathbb{R}^n$

$$\mathrm{Aff}[C] = \{\theta_1 x_1 + \ldots + \theta_k x_k \mid x_1, \ldots, x_k \in C, \sum_{i=1}^{k} \theta_i = 1\}$$

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:

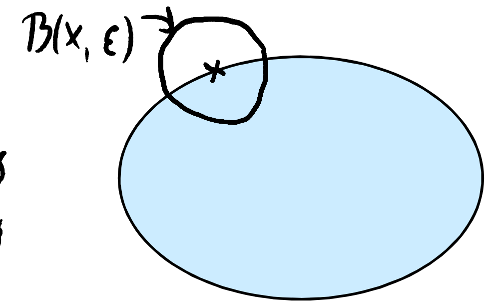
$$C = \{x \mid Ax = b\} \quad A \in \mathbb{R}^{m \times n}, \; b \in \mathbb{R}^{n}$$

Proof:

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

$$\theta \in \mathbb{R}$$

Boundaries



Definition [x on boundary of C (∂C)]:

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

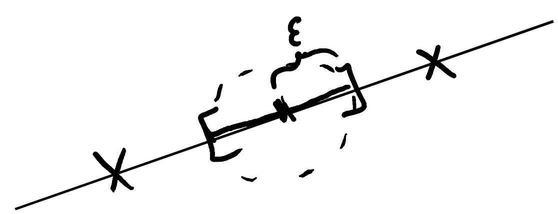
& $B(x,\varepsilon) \cap C^c \neq \emptyset$

Definition [x in interior of C]:

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

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

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



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

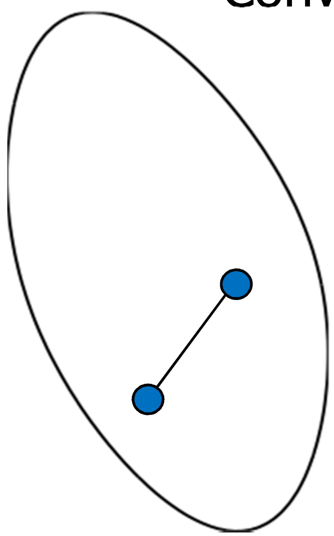
Definition [C closed]: $\partial C \subset C$

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

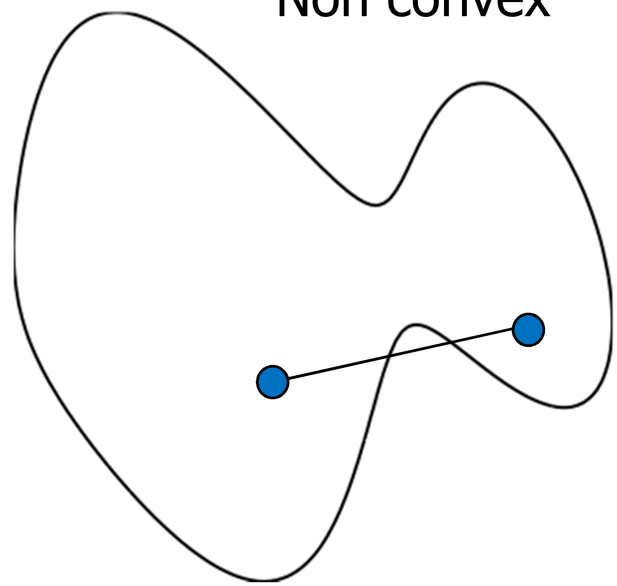
Definition [C compact]: CLOSED AND BOUNDED [IN $\mathbb{R}^n$]

Convex sets

Convex



Non convex



Definition [Convex set]:

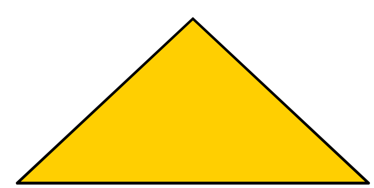
SET $C \subset \mathbb{R}^d$ IS CONVEX $\iff \forall x_1, x_2 \in C \;\; \forall \theta \in [0,1]$

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

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

Examples of Convex Sets

- Euclidian balls:

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

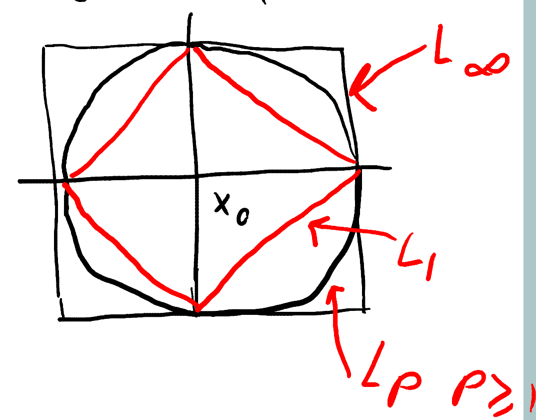
- L_p balls, $p \geq 1$

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

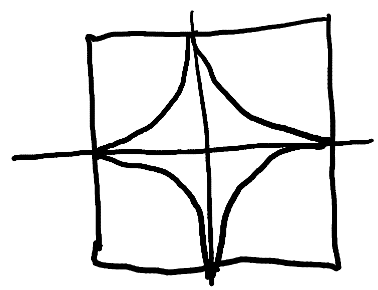
$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}$$

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

CONVEX



- L_p balls $0<p<1$



NOT CONVEX!

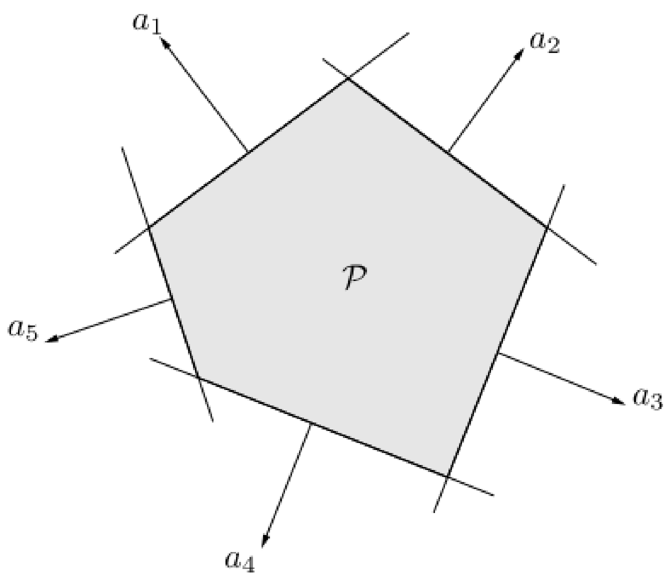
Examples of Convex Sets

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

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

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

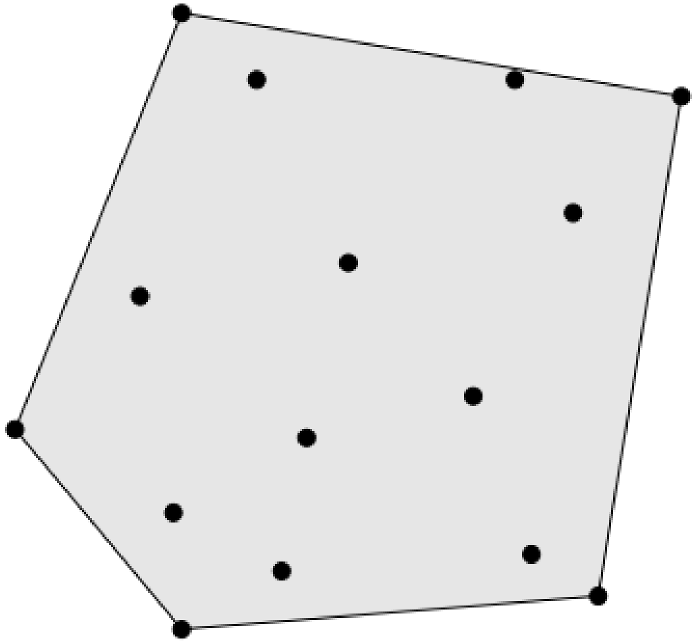
- ❑ Polytope: bounded polyhedron

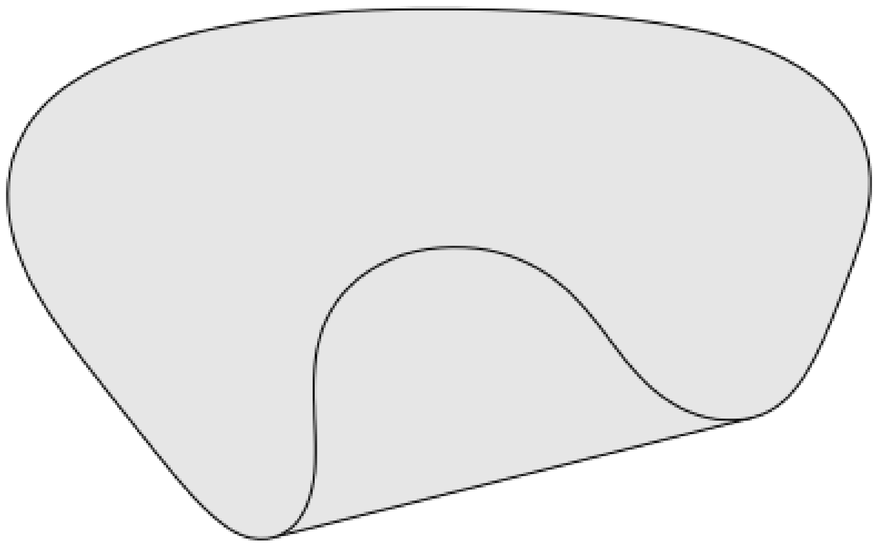


Intersection of halfspaces

& hyperplanes

Convex hull





Definition [Convex hull]:

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

Convex hull

Convex hull properties:

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

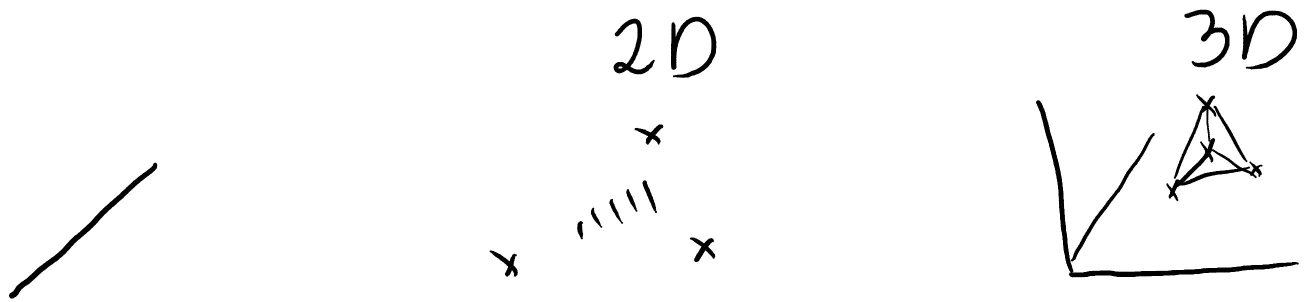
CONV[C] IS CONVEX

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

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

Examples of Convex Sets

- Simplex:



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 $\Rightarrow$ $\underbrace{\int_C p(x)\, x\, dx}_{\text{ASSUMING THIS INTEGRAL EXIST}} \in C$

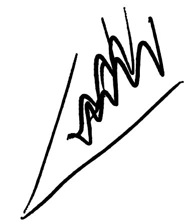
C CONVEX SET

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

Cones

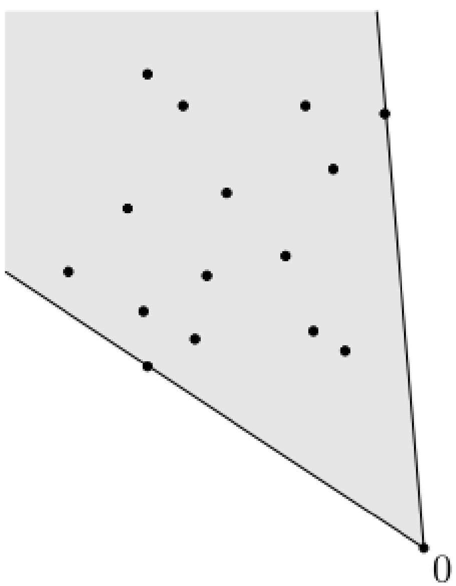
Definition [Cone]:

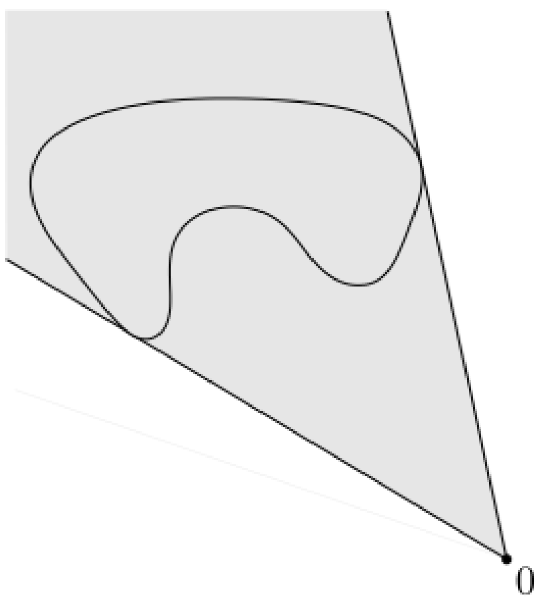
C is cone $\iff \left\{ \begin{array}{l} x \in C \\ \theta \geq 0 \end{array} \implies \theta x \in C \right\}$



Definition [Convex Cone]: cone & convex

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





Example: PSD matrices

Definition [Positive semi definite matrix]:

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

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:

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

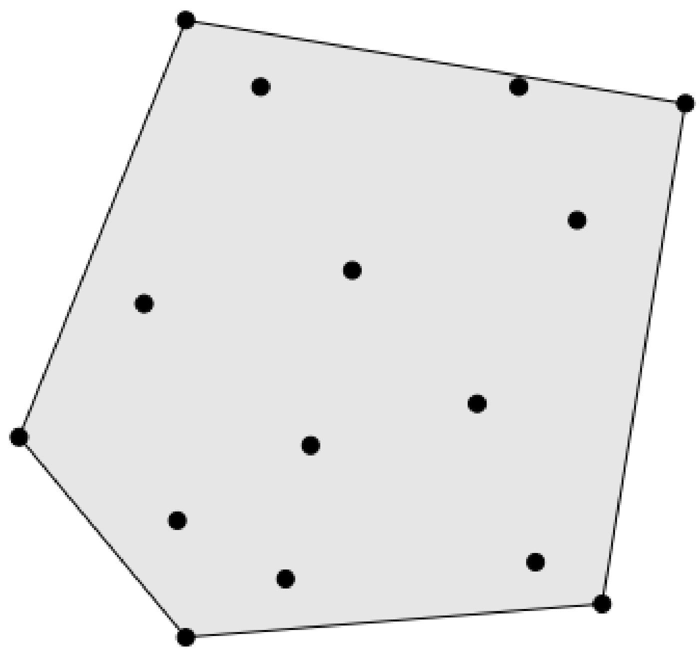
Proof:

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

Convex set representation with convex hull

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

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

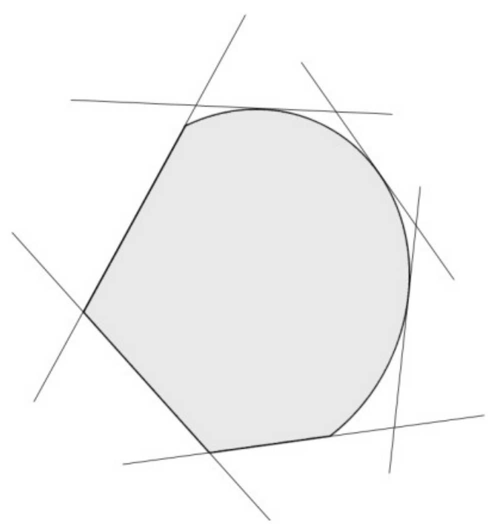


Dual representation

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

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



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

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

Convex Optimization
CMU-10725

5. Convexity Part II

Barnabás Póczos & Ryan Tibshirani

MACHINE LEARNING DEPARTMENT
Carnegie Mellon.
School of Computer Science

Conevex Hull

Definition 1

$$\text{conv}[C] = \left\{ \bigcap_{\alpha} S_\alpha \;\middle|\; S_\alpha \subset \mathbb{R}^n,\; C \subset S_\alpha,\; S_\alpha \text{ convex} \right\}$$

Note:

- If C is a finite set, then this is closed polyhedron.
- If C contains infinite many points, then this can be open, closed, or none of them

Theorem [Definition 2, Primal representation]

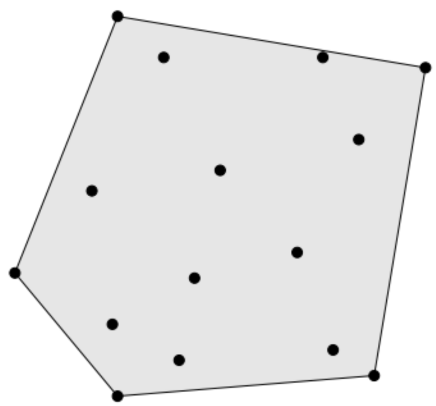
$$\text{conv}[C] = \left\{ \theta_1 x_1 + \ldots + \theta_k x_k \;\middle|\; x_i \in C,\; \theta_i \geq 0 \;\forall i = 1 \ldots k,\; \sum_{i=1}^{k} \theta_i = 1,\; k \in \mathbb{Z}_+ \right\}$$

A closed convex set is the intersection of all the closed half spaces containing S

Convex set representation with convex hull

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

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



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

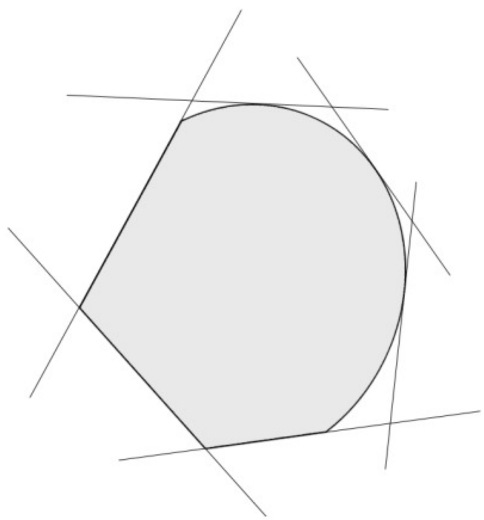
Convex hull = convex combination of possibly infinite many points in the set.

Dual representation

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

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

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



A closed convex set is the intersection of all the closed half spaces containing S

Convexity-preserving set operations

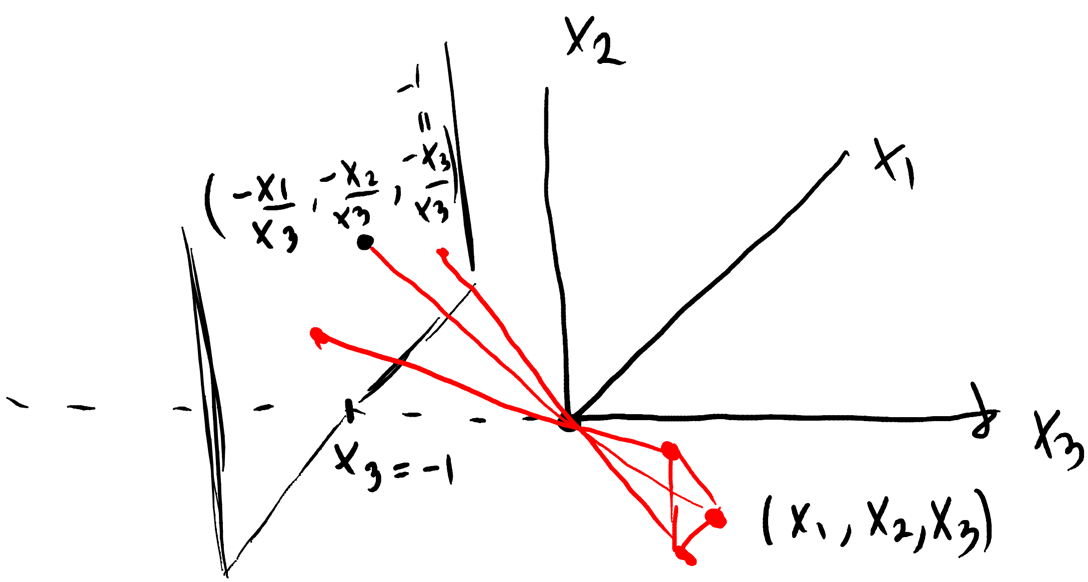
- Perspective projection (pinhole camera)

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)

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

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

Theorem: [Image of Linear fractional function]

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

Convexity-preserving set operations

Application: [Conditional probabilities]

$$C = \{ X = (p_{11}, p_{12}, \ldots, p_{mn}) \in [0,1]^{m \times n} \}$$

$U \in \{1, 2, \ldots, m\}$ DISCRETE

$V \in \{1, 2, \ldots, n\}$ RANDOM VARIABLES

$$p_{ij} = \mathrm{PROB}(U = i, V = j)$$

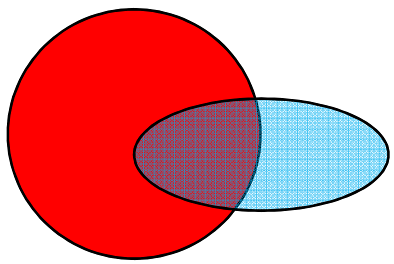
$$f_{ij} = \frac{p_{ij}}{\sum_{k=1}^{n} p_{kj}} = \mathrm{PROB}(U = i \mid V = j)$$

$$D = \{ y = (f_{11}, f_{12}, \ldots, f_{mn}) \in [0,1]^{m \times n} \}$$

IF C IS CONVEX $\Rightarrow$ D IS CONVEX

Convexity-preserving set operations

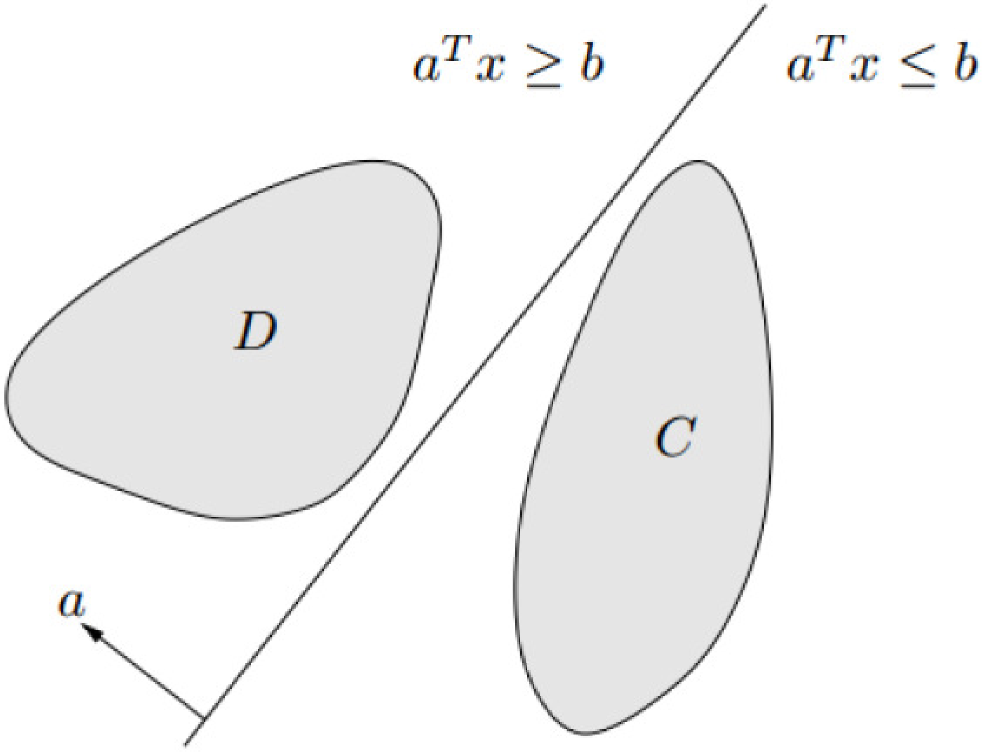
- Union doesn't preserve convexity



Separating hyperplane thm

Theorem: [Separating hyperplane theorem]

$$\left.\begin{array}{l} C, D \text{ convex sets} \\ C \cap D = \emptyset \end{array}\right\} \Rightarrow \exists\, a \in \mathbb{R}^n, \ b \in \mathbb{R} \text{ s.t. } \begin{array}{l} \forall x \in C \ \ a^T x \le b \\ \forall y \in D \ \ a^T y \ge b \end{array}$$



Separating hyperplane thm

Definition: [Strong separation]

$$a^T [C_1 + B(0,\varepsilon)] > b$$
$$a^T [C_2 + B(0,\varepsilon)] < b$$

Definition: [Proper separation],

IT'S NOT THE CASE THAT BOTH

$$C_1 \subseteq \{x : a^T x = b\}$$
$$C_2 \subseteq \{x : a^T x = b\}$$

Definition: [Strict separation]

$$a^T x > b \quad x \in C_1$$
$$a^T x < b \quad x \in C_2$$

It "strictly separates" them if neither one touches the hyperplane.

Separating hyperplane thm

Theorem: [Strong separation theorem]

C_1, C_2 NONEMPTY CONVEX SET IN $\mathbb{R}^n$

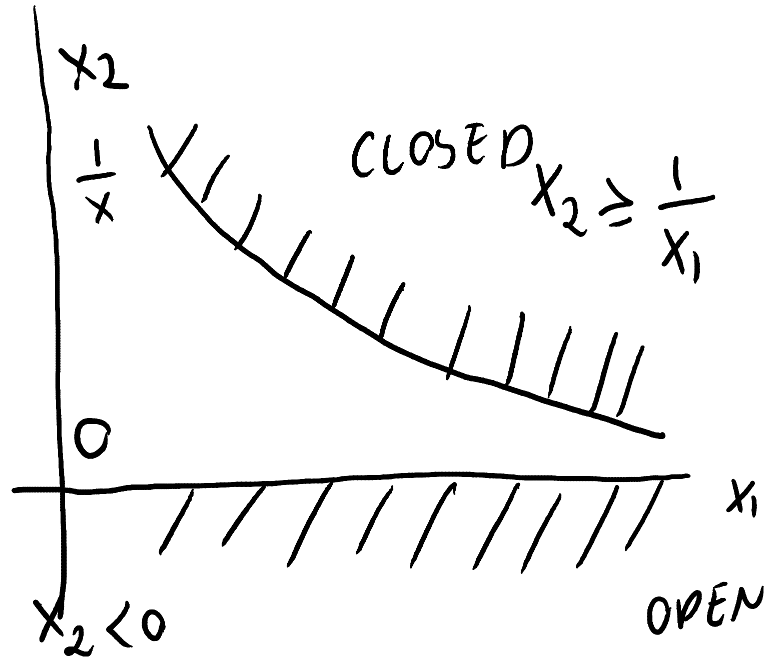
$\mathrm{cl}\, C_1 \cap \mathrm{cl}\, C_2 = \emptyset$

EITHER C_1 OR C_2 IS BOUNDED

$\Rightarrow$ $\exists$ HYPERPLANE SEPARATING C_1 AND C_2 STRONGLY

Counterexample:

Why do we need at least one bounded set?



$X_2 = 0$ SEPARATING HYPERPLANE

Separating hyperplane thm II

Theorem: [Strong separation theorem II]

C_1, C_2 ARE NONEMPTY CONVEX SETS

$C_1 \subseteq \mathbb{R}^n$, $C_2 \subseteq \mathbb{R}^n$

$\exists$ HYPERPLANE SEPARATING C_1 AND C_2 STRONGLY

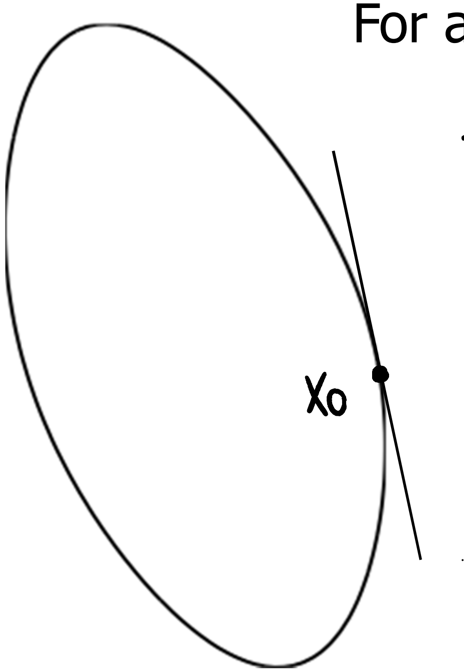
$$\iff \inf_{\substack{x_1 \in C_1 \\ x_2 \in C_2}} \{ |x_1 - x_2| \} > 0$$

$$\iff \mathrm{dist}(C_1, C_2) > 0$$

$$\iff 0 \notin \mathrm{cl}(C_1 - C_2)$$

Supporting hyperplane thm

Theorem: [Supporting hyperplane theorem]



For any point x_0 on the boundary of convex C

$\exists$ HYPERPLANE $\{x \mid a^T x = b\}$, $a \neq 0$

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

$a^T x_0 = b$

Theorem: [Partial converse of the supporting hyperplane theorem]

- SET C IS CLOSED
- $\mathrm{int}\, C \neq \emptyset$
- $x_0 \in \mathrm{bd}\, C \Rightarrow \exists$ SUPPORTING HYPERPLANE OF C

$\Rightarrow$ C IS CONVEX

Proving a set convex

- Use definition directly

- Represent as convex hull

- Represent as the intersection of halfspaces

- Supporting hyperplane partial converse:
  - C closed, nonempty interior, has supporting hyperplane at all boundary points $\Rightarrow$ C convex

- Build C up from simpler sets using convexity-preserving operations

Convex functions

Convex functions

Definition [convex function]:

A FUNCTION $f : \mathbb{R}^n \to \mathbb{R}$ IS CONVEX IF

- DOM f IS A CONVEX SET
- $f(\theta x + (1-\theta) y) \leq \theta f(x) + (1-\theta) f(y)$

$\forall x, y \in \text{DOM} f$

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

Definition [strictly convex function]:

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

$\forall x \neq y$

$\forall 0 < \theta < 1$

FOR EXAMPLE x^4 IS STRICTLY CONVEX

Concave functions

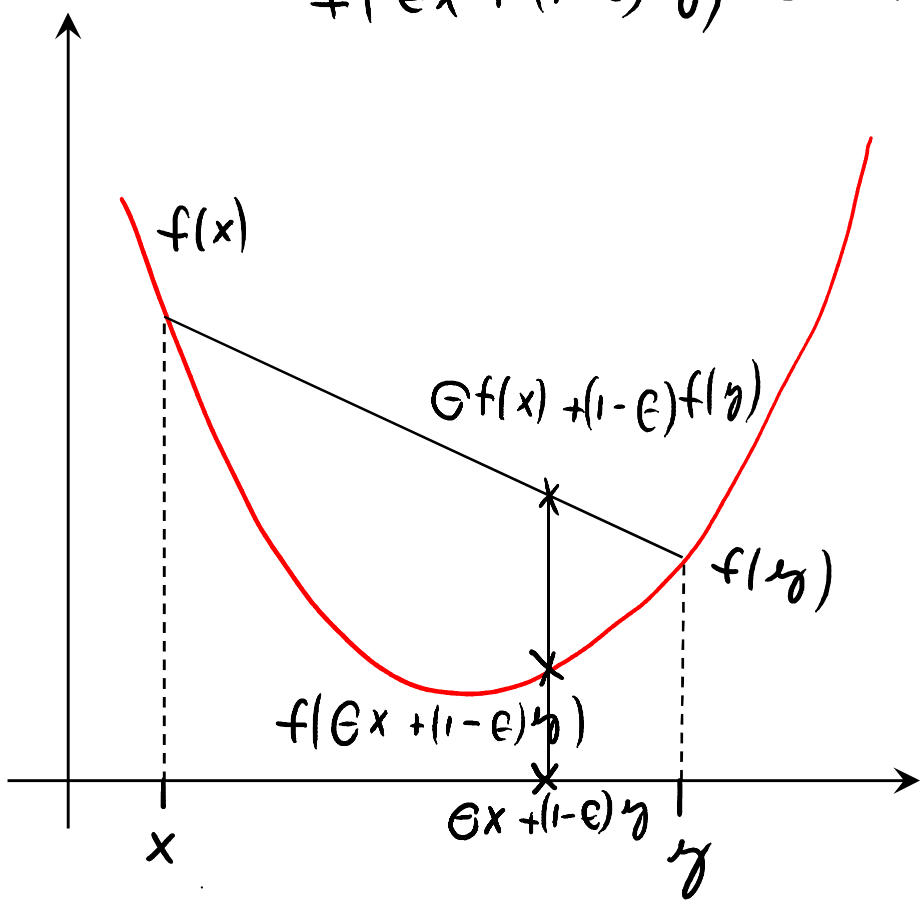
Definition [concave function]:

-f is convex

Convex functions

Geometric interpretation

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



Strongly convexity

Definition:[m-strongly convex function (m>0)

$$(\nabla f(x) - \nabla f(y))^T (x - y) \geq m\|x - y\|_2^2$$

An equivalent condition:

$$f(y) \geq f(x) + \nabla f(x)^T (y - x) + \frac{m}{2}\|y - x\|_2^2$$

Without gradient: $\forall t \in [0, 1]$

$$f(tx + (1 - t)y) \leq tf(x) + (1 - t)f(y) - \frac{1}{2}mt(1 - t)\|x - y\|_2^2$$

With Hessian:

$\nabla^2 f(x) \succeq mI$ for all x in the domain

A strongly convex function is also strictly convex, but not vice-versa.

Examples

$f(x) = x^4$: convex, strictly convex, not strongly convex.

$f(x) = |x|$: convex, not strictly convex.

Examples: Convex functions

Convex

- $|x|^p \quad p \geq 1$ ON $\mathbb{R}$
- $f(x) = \max(x_1, \ldots, x_n)$ ON $\mathbb{R}^n$
- ANY NORM

Concave

- GEOMETRIC MEAN $f(x) = \left(\prod_{i=1}^{n} x_i\right)^{1/n}$ CONCAVE ON $\mathbb{R}^n_{++}$
- $\log \det(X)$ IS CONCAVE ON S^n_{++}

Extended reals

We can extend f from dom f to R^n without changing its convexity

LET $\tilde{f} : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$

$$\tilde{f}(x) = \begin{cases} f(x) & x \in \text{DOM } f \\ \infty & x \notin \text{DOM } f \end{cases}$$

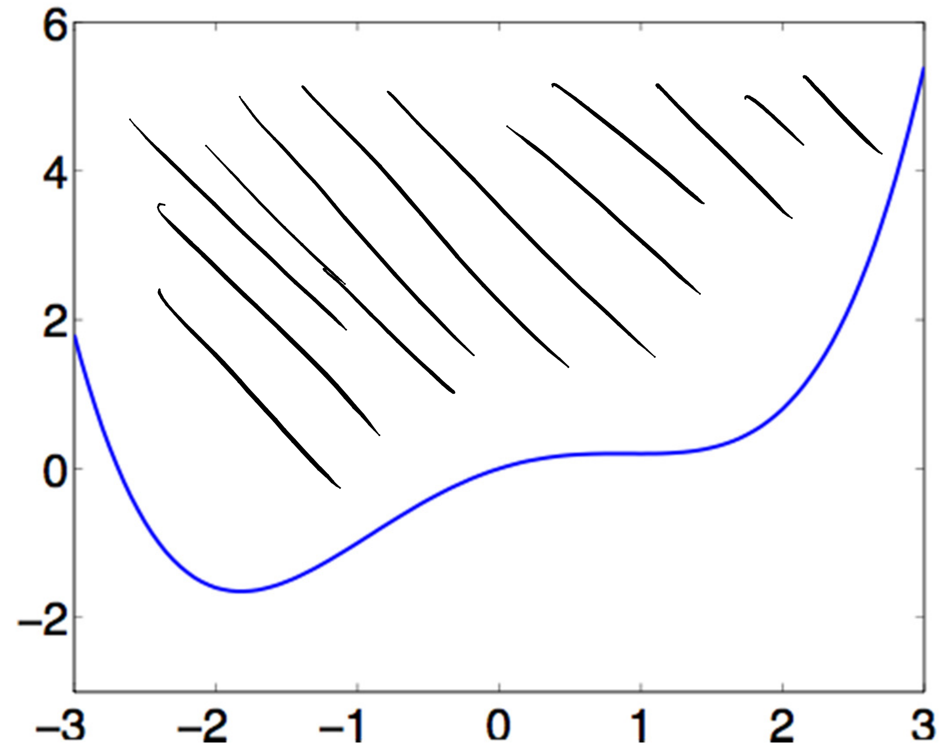
Theorem:

f IS CONVEX $\iff$ $\tilde{f}$ IS CONVEX

$\iff \tilde{f}(\theta x + (1-\theta)y) \leq \theta \tilde{f}(x) + (1-\theta)\tilde{f}(y)$

Epigraph

Definition [epigraph]: $\mathrm{EPI}(f) = \{(x,t) : x \in \mathrm{DOM}\, f,\ t \geq f(x)\}$



Theorem [convexity of the epigraph]:

$f: \mathbb{R}^n \to \mathbb{R}$ IS CONVEX $\Longleftrightarrow$ $\mathrm{EPI}(f)$ IS CONVEX

Convex Function Properties

0th order characterization

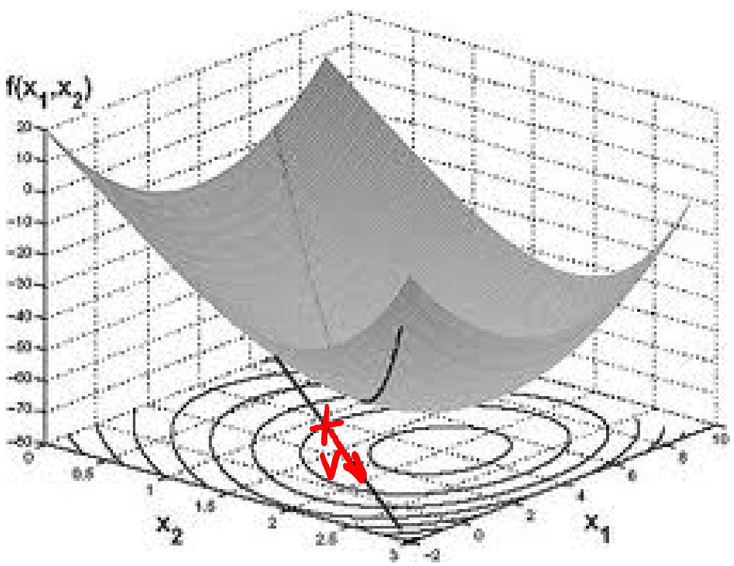
f convex $\iff g(t) = f(x+tv)$ is convex

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

on $\text{dom } g = \{t \mid x+tv \in \text{dom } f\}$

$\forall x \in \text{dom } f$ and $\forall v \in \mathbb{R}^n$

This is useful, because we only need to check the convexity of 1D functions.



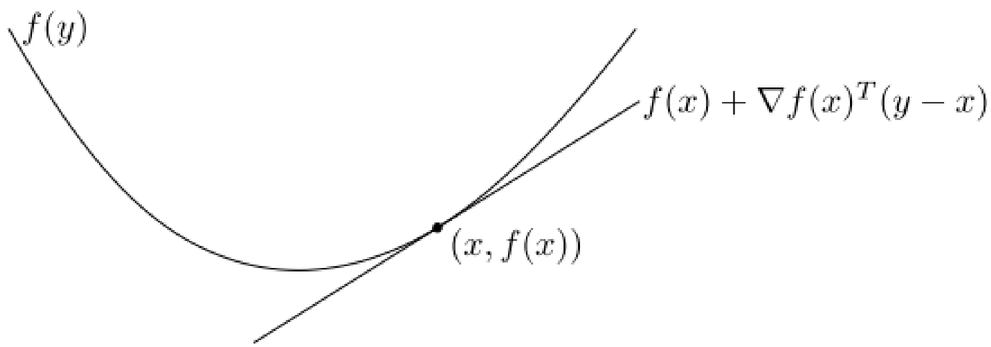
Graph courtesy of Prof. Robert Freund

Convex Function Properties

1st order characterization

- LET f BE A DIFFERENTIABLE FUNCTION
- DOM f OPEN, CONVEX

f IS CONVEX $\iff f(y) \geq f(x) + \nabla f(x)^T (y-x) \quad \forall y$



The 1st oder Taylor approximation is a global underestimator of f.

Corollary: $\left. \begin{array}{l} f \text{ CONVEX} \\ \nabla f(x) = 0 \end{array} \right\} \Rightarrow f(y) \geq f(x) \quad \forall y$

x IS GLOBAL OPTIMUM

Convex Function Properties

2nd order characterization

- LET f BE TWICE DIFFERENTIABLE
- DOM f OPEN

$$f \text{ IS CONVEX} \iff \nabla^2 f(x) \succeq 0 \quad \forall x \in \operatorname{dom} f$$

Lemma

IF $\nabla^2 f(x) \succ 0 \ \forall x \in \operatorname{dom} f \implies f$ IS STRICTLY CONVEX

$\not\Leftarrow$ FOR EXAMPLE X^4

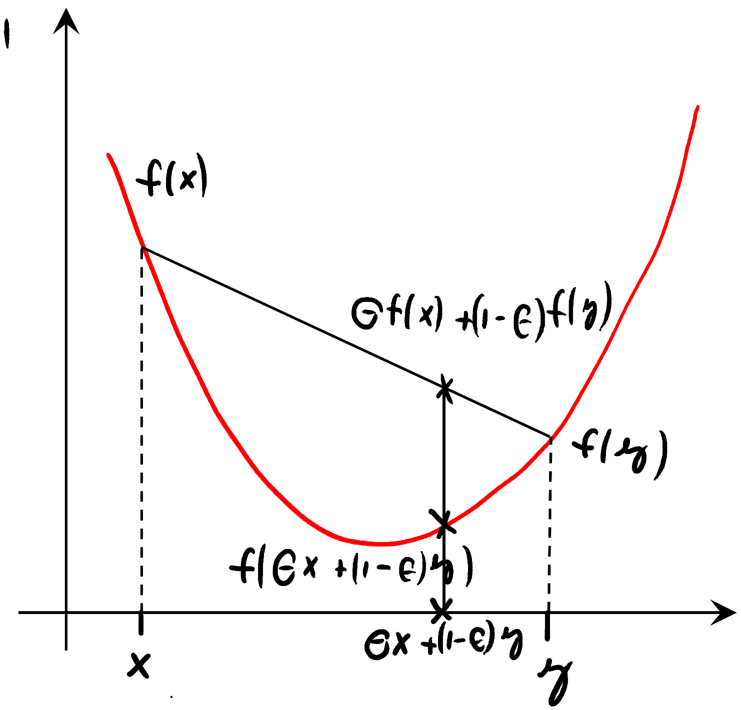
Jensen's inequality

Theorem

$f: \mathbb{R}^n \to (-\infty, \infty]$

f is convex $\iff f(\lambda_1 x_1 + \dots + \lambda_m x_m) \le \lambda_1 f(x_1) + \dots + \lambda_m f(x_m)$

$\forall\ \lambda_1 \geq 0 \dots \lambda_m \geq 0,\ \sum_{i=1}^{m} \lambda_i = 1$



f is convex $\Rightarrow f(\mathbb{E}X) \le \mathbb{E} f(X)$

$X \sim p$

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

Proving a function convex

- Use definition directly
- Prove that epigraph is convex via set methods
- 0^{th}, 1^{st}, 2^{nd} order convexity properties
- Construct f from simpler convex fns using convexity-preserving ops

Convexity-preserving fn ops

Nonnegative weighted sum

If f_1, f_2 cvx, $w_i \geq 0 \Rightarrow h(x) = w_1 f_1(x) + w_2 f_2(x)$ cvx

Pointwise max/sup

If f, g cvx, $\Rightarrow m(x) = \max\{f(x), g(x)\}$ cvx

Extension of pointwise max/sup

If $f(x, y)$ is convex in x for each y

$\Rightarrow g(x) = \sup_{y \in C} f(x, y)$ is convex in x,

provided $g(x) > -\infty$ for some x.

Convexity-preserving fn ops

Affine map

If $f : \mathbb{R}^n \to \mathbb{R}$ is cvx, $\Rightarrow g(x) = f(Ax + b)$ is cvx, where $A \in \mathbb{R}^{n \times m}$, $b \in \mathbb{R}^n$.

Composition

If f, g are cvx, and g is non-decreasing,
$\Rightarrow h(x) = g(f(x))$ is cvx.

If f is concave and g is cvx and non-increasing,
$\Rightarrow h(x) = g(f(x))$ is cvx.

Perspective map

If $f(x)$ is convex, $\Rightarrow g(x, t) = tf(x/t)$ is convex.

Summary

- **Convex sets**
  - Representation:
    - convex hull, intersect hyperplanes
  - supporting, separating hyperplanes
  - operations that preserve convexity

- **Convex functions**
  - epigraph
  - 0 orders, 1st order, 2nd order conditions
  - operations that preserve convexity