# Convex Sets

Optimization - 10725
Carlos Guestrin
Carnegie Mellon University

February 25th, 2008




# Announcements

- Class project:
  - Opportunity to explore interesting optimization problem of your choice.
  - May involve optimization in some problem in ML, AI or other domain of your interest, or to implement and evaluate core optimization techniques.
    - Some ideas in the website.
  - All projects must have an implementation component, though theoretical aspects may also be explored.
  - You should evaluate your approach, preferably on real-world data.
  - It will be fun!!! :)
- Fine print:
  - Class project must be about new things you have done this semester; you can't use results you have developed in previous semesters.
  - Individual or groups of 2 students.
  - Deliverables:
    - Brief project proposal (1 page) by **March 5th** in class.
    - Midway progress report (5 pages) describing the results of your first experiments by **April 9th** in class, worth 20% of the project grade.
    - Poster for class poster session on **May 1st, 3-6pm** in the NSH Atrium, worth 20%.
    - Write up (8 pages maximum in NIPS format, including references; this page limit is strict), due **May 5th by 3pm** by email, worth 60% of the project grade.

## Convex optimization v. Nonlinear optimization

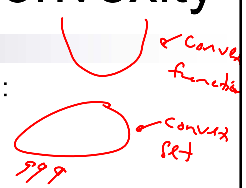
- Linear optimization problems
  - Linear objective, linear constraints
  - Efficient solutions!
- Nonlinear optimization
  - Either nonlinear constraints or objective
  - You will often hear: "problem is nonlinear, no hope to solve it… must use local search, simulated annealing,…"
- Convex optimization
  - Many nonlinear objectives/constraints are convex
  - Efficient solutions
- Real question: "convex v. non-convex?"
  - Not "linear v. nonlinear?"
- Even if problem is non-convex, convexity is useful:
  - Convex relaxations of non-convex problems may have theoretical guarantees
  - Can always obtain convex lower bound to non-convex problem
    - Duality (always) and relaxation (often)
  - Can provide good starting point for local search




## Outline to learning about convexity

- General definition of a convex optimization problem:

$$\min_x f(x) \leftarrow \text{convex function}$$
$$x \in C \leftarrow \text{convex set}$$



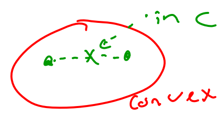

- Equivalent problem:

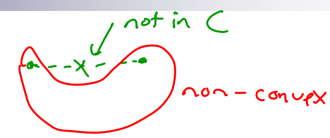
$$\min_x f(x)$$
$$g_i(x) \leq b_i \quad \forall i \qquad g_i \leftarrow \text{convex function}$$

- How we'll learn about these problems:
  1. Convex sets
  2. Convex functions
  3. Convex optimization problems
  4. Duality and convexity
  5. Algorithms for optimizing convex problems
- Applications will be discussed along the way
- Today: characterizing convex sets and some interesting examples

# Definitions of convex sets

- Convex v. Non-convex sets



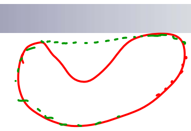




- Line segment definition: $\forall x_1, x_2 \in C, \forall \theta \in [0,1]$, $z = \theta x_1 + (1-\theta) x_2 \Rightarrow z \in C$
- Convex combination definition: $\theta_1, \ldots \theta_k \geq 0$, $\sum_i \theta_i = 1$, if $x_1 \ldots x_k \in C \Rightarrow \sum_i \theta_i x_i \in C$
- Probabilistic interpretation:
  - If $C \subseteq R^n$ is convex
  - Define a probability distribution $p(x)$: $p(x) \geq 0$, $\int_{x \in C} p(x)\,dx = 1$
  - Then $E[x] \in C$



# General convex hull

- Given some set $C$



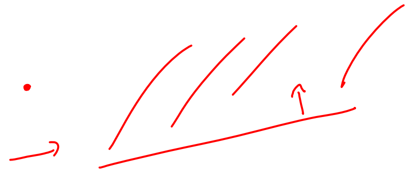
- Convex hull of $C$, **conv** $C$

$$\text{conv } C = \{x \mid x = \sum_i \theta_i x_i,\ x_i \in C,\ \theta_i \geq 0, \sum \theta_i = 1\}$$

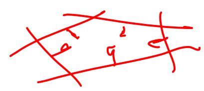
- Properties of convex hull:
  - Idempotency: $C \in$ convex $\text{conv } C = C$, $\text{conv } C = \text{conv conv } C$
  - Convexity:
- Usefulness: obtain a lower bound on non-convex problem $\min_x f(x)$ s.t. $x \in C$ $\rightarrow$ $\min_x f(x)$, $x \in \text{conv } C$

# Examples of convex sets we have already seen…

- $R^n$
- point
- half space



- polyhedron



- line
- line segment
- linear subspace $x \mid Ax = b$



# First non-linear example: Euclidean balls and Ellipsoids

- $B(x_c,r)$ - ball centered at $x_c$ centered at r:

$$B(x_c,r) = \{x \mid \|x-x_c\|_2 \leq r\} = \{x \mid \sqrt{(x-x_c)^T(x-x_c)} \leq r\}$$

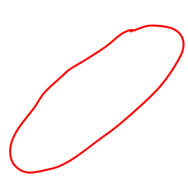
- Convexity:

$$x_1, x_2 \in B(x_c,r), \quad \theta x_1 + (1-\theta)x_2 \in B(x_c,r)$$

$$\|\theta_1 x_1 + (1-\theta)x_2 - x_c\|_2 = \|\theta x_1 + (1-\theta)x_2 - \theta x_c - (1-\theta)x_c\|_2$$
$$\leq \|\theta x_1 - \theta x_c\|_2 + \|(1-\theta)x_2 - (1-\theta)x_c\|_2$$
$$= \theta\|x_1 - x_2\|_2 + (1-\theta)\|x_2 - x_c\|_2 \leq r$$

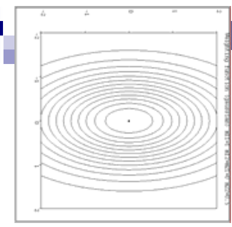
cool!!

- Ellipsoid:
  - $(x-x_c)^T \Sigma^{-1} (x-x_c) \leq 1$
  - $\Sigma$ is positive semidefinite

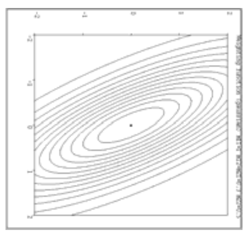


convex.

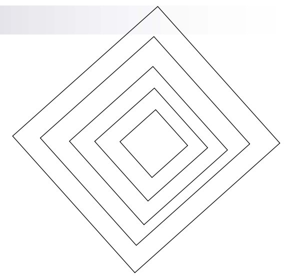
# Examples of Norm Balls



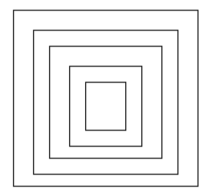
**Scaled Euclidian ($L_2$)**



**Mahalanobis**



**$L_1$ norm (absolute)**



**$L_\infty$ *(max) norm***




# Norm balls

- Convexity of norm balls
  - Properties of norms:
    - Scaling
    - Triangle inequality
- Norm balls are extremely important in ML
  -
- What about achieving a norm with equality?

# Cones

- Set *C* is a cone if set is invariant to non-negative scaling

- If the cone is convex, we call it:
  - extremely important in ML (as we'll see)

- A cool cone: The ice cream cone
  - a.k.a. second order cone




# Positive semidefinite cone

- Positive semidefinite matrices:

- Positive semidefinite cone:

QuickTime™ and a
TIFF (LZW) decompressor
are needed to see this picture.

- Alternate definition: Eigenvalues

- Convexity:

$$X = \begin{bmatrix} x & y \\ y & z \end{bmatrix}$$

- Examples in ML:
  - 
  - 
- A fundamental convex set
  - Useful in a huge number of applications
  - Basis for very cool approximation algorithms
  - Generalizes pretty many "named" convex optimization problems

# Operations that preserve convexity 1: Intersection

- Intersection of convex sets is convex
- Examples:
  - Polyhedron
  - Robust linear regression
  - Positive semidefinite cone




# Operations that preserve convexity 2: Affine functions

- Affine function: f(x) = Ax + b
- Set *S* is convex
  - Image of *S* under f is convex
- Translation:
- Scaling:
- General affine transformation:
- Why is ellipsoid convex?
  - $(x-x_c)^T \Sigma^{-1}(x-x_c) \leq 1$
  - $\Sigma$ is positive semidefinite

## Operations that preserve convexity 3: Linear-fractional functions

- Linear fractional functions:
  - 
  - Closely related to perspective projections (useful in computer vision)
- Given convex set *C*, image according to linear fractional function:
- Example:




## Separating hyperplane theorem

- **Theorem**: Every two non-intersecting convex sets *C* and *D* have a separating hyperplane:
- Intuition of proof (for special case)
  - Minimum distance between sets:
  - If minimum is achieved in the sets (e.g., both sets closed, and one is bounded), then

# Supporting hyperplane

- General definition: Some set $C \subseteq R^n$
  - Point $x_0$ on boundary
    - Boundary is the closure of the set minus its interior
  - Supporting hyperplane:
    - Geometrically: a tangent at $x_0$
    - Half-space contains *C*:

- **Theorem**: for any non-empty convex set *C*, and any point $x_0$ in the boundary of *C*, there exists (at least one) supporting hyperplane at $x_0$

- (One) **Converse**: If set *C* is closed with non-empty interior, and there is a supporting hyperplane at every boundary point, then *C* is convex




# What you need to know

- Definitions of convex sets
  - Main examples of convex sets
- Proving a set is convex
- Operations that preserve convexity
  - There are many many many other operations that preserve convexity
    - See book for several more examples
- Separating and supporting hyperplanes