Convex Sets (cont.) Convex Functions

Optimization - 10725

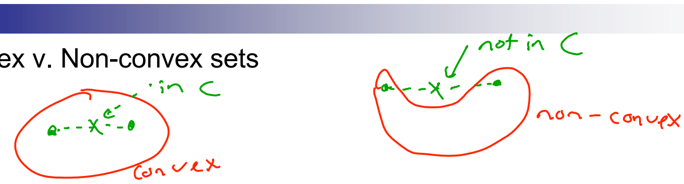
Carlos Guestrin

Carnegie Mellon University

February 27th, 2008

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:

  $\theta_1, \dots, \theta_k \geq 0$, if $x_1 \dots x_k \in C \Rightarrow \sum_i \theta_i x_i \in C$

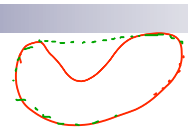
  $\sum_i \theta_i = 1$

- Probabilistic interpretation:
  - If $C \subseteq R^n$ is convex
  - Define a(ny) probability distribution $p(x)$: $p(x) \geq 0$, $\int_{x \in C} p(x)dx = 1$
  - Then $E[X] \in C$

see slides for detail ? $E[x]$ exist → fancy math

General convex hull

- Given some set *C*



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

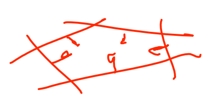
$\mathrm{conv}\, C = \{x \mid x = \sum \theta_i x_i,\ x_i \in C,\ \theta_i \geq 0, \sum \theta_i = 1\}$

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

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:

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

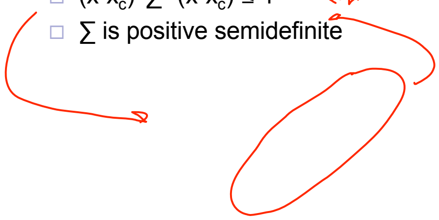
- Convexity:

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

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

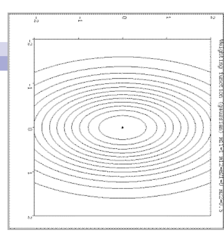
cool!! :)

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

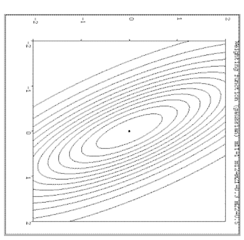


convex.

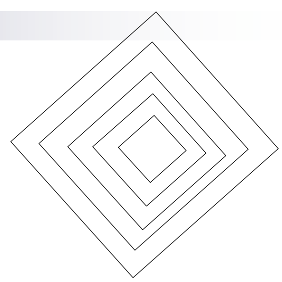
Examples of Norm Balls



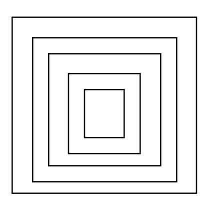
Scaled Euclidian (L_2)



Mahalanobis



L_1 norm (absolute)



L_∞ *(max) norm*

Norm balls

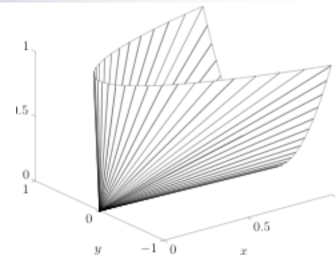
- Convexity of norm balls
  - Properties of norms:
    - Scaling
    - Triangle inequality
- Norm balls are extremely important in ML
  -
- What about achieving a norm with equality?

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:
- Alternate definition: Eigenvalues
- Convexity:
- 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



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

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?
  - $(x-x_c)^T\Sigma^{-1}(x-x_c) \leq 1$
  - Σ 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

Convex Functions

- Function $f: R^n \rightarrow R$ is convex if
  - Domain is convex
  -
- Generalization: Jensen's inequality:
- Strictly convex function:

Concave functions

- Function f is concave if
  -
  -
- Strictly concave:
- We will be able to optimize:

Proving convexity for a very simple example

- $f(x)=x^2$

First order condition

- If f is differentiable in all **dom** f
- Then f convex if and only if **dom** f is convex and

Second order condition (1D f)

- If f is twice differentiable in **dom** f
- Then f convex if and only if **dom** f is convex and
- Note 1: Strictly convex if:
- Note 2: **dom** f must be convex
  - $f(x)=1/x^2$
  - **dom** f = $\{x \in R | x \neq 0\}$

Second order condition (general case)

- If f is twice differentiable in **dom** f
- Then f convex if and only if **dom** f is convex and
- Note 1: Strictly convex if:

Quadratic programming

- $f(x) = (1/2)\ x^T A x + b^T x + c$
- Convex if:
- Strictly convex if:
- Concave if:
- Strictly concave if:

Simple examples

- Exponentiation: e^{ax}
  - convex on R, any $a \in R$
- Powers: x^a on R_{++}
  - Convex for $a \leq 0$ or $a \geq 1$
  - Concave for $0 \leq a \leq 1$
- Logarithm: log x
  - Concave on R_{++}
- Entropy: -x log x
  - Concave on R_+
  - (0 log 0 = 0)

A few important examples for ML

- Every norm on R^n is convex
- Log-sum-exp:
  - Convex in R^n
- Log-det:
  - Convex in S^n_{++}

Extended-value extensions

- Convex function f over convex **dom** f
- Extended-value extension:
- Still convex:
- Very nice for notation, e.g.,
  - Minimization:
  - Sum:
    - f_1 over convex **dom** f_1
    - f_2 over convex **dom** f_2

Epigraph

- Graph of a function $f:R^n \rightarrow R$
  - $\{(x,t) \mid x \in$ **dom** $f, f(x)=t\}$
- Epigraph:
  - **epi** f =
- **Theorem**: f is convex if and only if

Support of a convex set and epigraph

- If f is convex & differentiable
  - $f(x) \geq f(x_0) + \nabla f(x_0)(x - x_0)$
- For $(x,t) \in$ **epi** f, $t \geq f(x)$, thus:
  -
- Rewriting:

$$(x,t) \in \mathbf{epi} f \Rightarrow \left[\begin{array}{c} \nabla f(x_0) \\ -1 \end{array} \right]^T \left(\left[\begin{array}{c} x \\ t \end{array} \right] - \left[\begin{array}{c} x_0 \\ f(x_0) \end{array} \right] \right) \leq 0$$

- Thus, if convex set is defined by epigraph of convex function
  - Obtain support of set by gradient!!
  - If f is not differentiable