EE445 Mod4-Lec1: Convexity and Convex Sets

References: [OM] Chapter 8

Topics for Module 4

- 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

$$\begin{array}{ll} \text{minimize}_x & f_0(x) \\ \text{subject to} & f_i(x) \leq b_i, \quad i = 1, \ldots, m, \end{array}$$

with components

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

Optimization: Applications

Many applications:

- Data fitting and regression
- Classification
- Image processing
- Portfolio optimization
- Recommender systems
- Optimal control
- 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

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."[1]

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

An important class: **convex optimization** problems

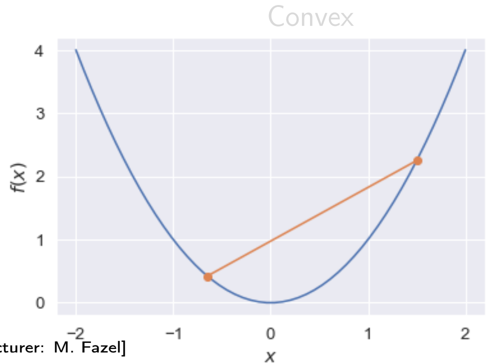
> "With only a bit of exaggeration, we can say that if you formulate a practical problem as a convex optimization problem, then you have solved the original problem."[1]

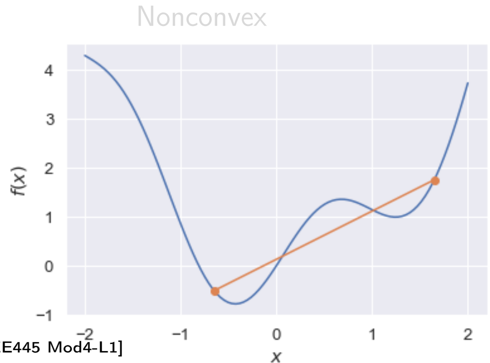
Convex Optimization

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 $0 \le \lambda \le 1$.





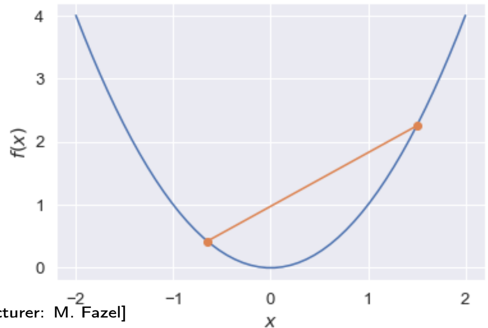
Convex Optimization

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

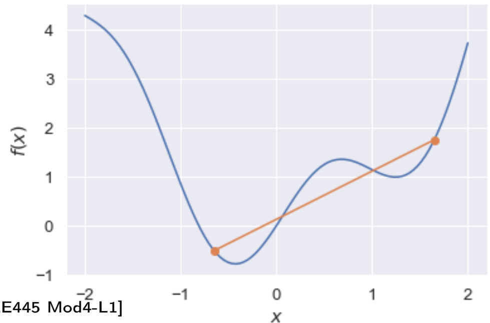
$$f_i\Big(\lambda x + (1-\lambda)y\Big) \leq \lambda f_i(x) + (1-\lambda)f_i(y)$$

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 $\{x \in \mathbf{R}^n \mid f_i(x) \leq 0\}$ 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 hulls, and give examples

Subspaces

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

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

Affine sets

$S \subseteq \mathbf{R}^n$ is *affine* if for $x, y \in S,\ \ \lambda, \mu \in \mathbf{R},\ \ \lambda + \mu = 1 \Longrightarrow \lambda x + \mu y \in S$
geometrically: $x, y \in S \Rightarrow$ line through $x, y \subseteq S$

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

$$x, y \in S, \;\; \lambda, \mu \geq 0, \;\; \lambda + \mu = 1 \Longrightarrow \lambda x + \mu y \in S$$

or equivalently,

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

geometrically: $x, y \in S \Rightarrow$ segment $[x, y] \subseteq S$
examples (one convex, two nonconvex sets)

Combinations and hulls

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

- *linear combination* of x_1, ..., 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:	$\mathbf{Cone}(S)$

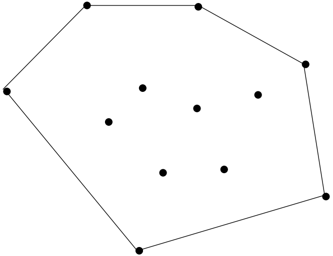
Convex combination and convex hull

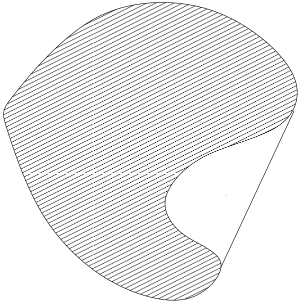
convex combination of x_1,..., x_k: any point x of the form

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

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

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





Hyperplanes and halfspaces

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

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

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

Euclidean balls and ellipsoids

(Euclidean) ball with center x_c and radius r:

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

ellipsoid: set of the form

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

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

Norm balls and 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

ℓ_p norms

ℓ_p norms on $\mathbf{R}^n$: for $p \geq 1$, $\|x\|_p = (\sum_i |x_i|^p)^{1/p}$,, 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 $\mathbf{R}^2$):

Polyhedra

solution set of finitely many linear inequalities and equalities

$$Ax \preceq b, \qquad Cx = d$$

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

polyhedron is intersection of finite number of halfspaces and hyperplanes

Operations that preserve convexity

practical methods for establishing convexity of a set C

1. apply definition

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

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.

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 $m = 2$:

Convexity preserved under *affine function*

suppose $f : \mathbf{R}^n \to \mathbf{R}^m$ is affine ($f(x) = Ax + b$ with $A \in \mathbf{R}^{m \times n}$, $b \in \mathbf{R}^m$)

- the image of a convex set under f is convex

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

examples

- scaling, translation
- projection