Convex Optimization

David Rosenberg

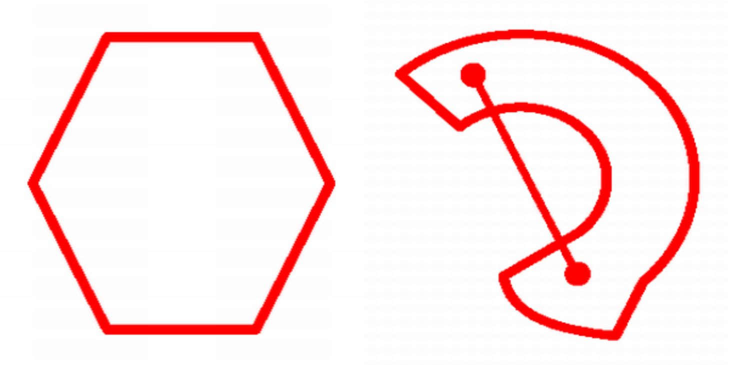
New York University

February 4, 2015

Convex Sets

Definition

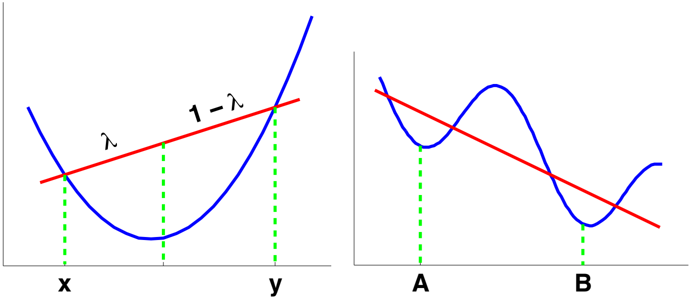
A set C is **convex** if the line segment between any two points in C lies in C.



Convex and Concave Functions

Definition

A function $f : \mathbf{R}^n \to \mathbf{R}$ is **convex** if the line segment connecting any two points on the graph of f lies above the graph. f is **concave** if $-f$ is convex.



Examples of Convex Functions on $\mathbf{R}$

Examples

- $x \mapsto e^{ax}$ is convex on $\mathbf{R}$ for all $a \in \mathbf{R}$
- $x \mapsto x^a$ is convex on $\mathbf{R}_{++}$ when $a \geqslant 1$ or $a \leqslant 0$ and concave for $0 \leqslant a \leqslant 1$
- $|x|^p$ for $p \geqslant 1$ is convex on $\mathbf{R}$
- $\log x$ is concave on $\mathbf{R}^{++}$
- $x \log x$ (either on $\mathbf{R}_{++}$ or on $\mathbf{R}_{+}$ if we define $0 \log 0 = 0$) is convex

Examples of Convex Functions on $\mathbf{R}^n$

Examples

- Every norm on $\mathbf{R}^n$ is convex
- Max: $(x_1, \ldots, x_n) \mapsto \max\{x_1 \ldots, x_n\}$ is convex on $\mathbf{R}^n$
- Log-Sum-Exp: $(x_1, \ldots, x_n) \mapsto \log\left(e^{x_1} + \cdots + e^{x_n}\right)$ is convex on $\mathbf{R}^n$.

Convex Functions and Optimization

Definition

A function f is **strictly convex** if the line segment connecting any two points on the graph of f lies **strictly** above the graph (excluding the endpoints).

Consequences for optimization:

- **convex:** if there is a local minimum, then it is a **global** minimum
- **strictly convex:** if there is a local minimum, then it is the **unique global** minumum

Convex Optimization Problem: Standard Form

Convex Optimization Problem: Standard Form

$$
\begin{array}{ll}
\text{minimize} & f_0(x) \\
\text{subject to} & f_i(x) \leqslant 0, \;\; i = 1, \ldots, m \\
& a_i^T x = b_i, \;\; i = 1, \ldots p
\end{array}
$$

where $f_0, \ldots, f_m$ are convex functions.
f_0 is called the **objective function.**
f_i are called the **inequality constraint functions.**

Convex Optimization Problem: More Terminology

- The set of points satisfying the constraints is called the **feasible set**.
- A point x in the feasible set is called a **feasible point.**
- The **optimal value** p^* of the problem is defined as

$$p^* = \inf\{f_0(x) \mid f_i(x) \leqslant 0, i = 1, \ldots, m,\ h_i(x) = 0,\ i = 1, \ldots, p\}.$$

- x^* is an **optimal point** (or a solution to the problem) if x^* is feasible and $f(x^*) = p^*$.

Convex Optimization Problem: Local Optimality

- We say that a feasible point x is **locally optimal** if there is an $R > 0$ such that x solves the following optimization problem:

$$\begin{array}{ll} \text{minimize} & f_0(z) \\ \text{subject to} & f_i(z) \leqslant 0, \;\; i = 1, \ldots, m \\ & h_i(z) = 0, \;\; i = 1, \ldots p \\ & \|z - x\|_2 \leqslant R \end{array}$$

  with optimization variable z.
  - Roughly speaking, this means x minimizes f_0 over nearby points in the feasible set.

Fact

A fundamental property of convex optimization problems is that any locally optimal point is also globally optimal.

Why Convex Optimization?

- Historically:
  - **Linear programs** (linear objectives & constraints) were the focus
  - **Nonlinear programs**: some easy, some hard
- Today:
  - Main distinction is between **convex** and **non-convex** problems
  - Convex problems are the ones we know how to solve efficiently
- Many techniques that are well understood for convex problems are applied to non-convex problems
  - e.g. SGD is routinely applied to neural networks

Your Reference for Convex Optimization

- Boyd and Vandenberghe (2004)
  - Very clearly written, but has a ton of detail for a first pass.
  - See my "Extreme Abridgement of Boyd and Vandenberghe".