Introduction to convex optimization

Prepared by Nia Hodges | October 9th, 2024

Washington University in St. Louis

"… in fact, the great watershed in optimization isn't between linearity and nonlinearity, but convexity and nonconvexity."

- R. Tyrrell Rockafellar, 1993

"… in fact, the great watershed in optimization isn't between linearity and nonlinearity, but convexity and nonconvexity."

- R. Tyrrell Rockafellar, 1993

Easy	Hard
Linear	Nonlinear

"… in fact, the great watershed in optimization isn't between linearity and nonlinearity, but convexity and nonconvexity."

- R. Tyrrell Rockafellar, 1993

Easy	Hard
Linear	Nonlinear
Convex	Nonconvex

Convexity generalizes the notion of linearity.

Initial optimization problem

$$\begin{array}{lll} \text{minimize} & f_0(x) & \\ \text{subject to} & f_i(x) \leq 0, & i = 1, \ldots, m \\ & g_i(x) = 0, & i = 1, \ldots, p \end{array}$$

- $x \in \mathbf{R}^n$ is (vector) variable to be chosen (n scalar variables $x_1, \ldots, x_n$)
- f_0 is the **objective function**, to be minimized
- $f_1, \ldots, f_m$ are the **inequality constraint functions**
- $g_1, \ldots, g_p$ are the **equality constraint functions**

[Source: Boyd and Vandenberghe]

Convex optimization

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m \\ & Ax = b \end{array}$$

- variable $x \in \mathbf{R}^n$
- equality constraints are linear
- $f_0, \ldots, f_m$ are **convex**: for $\theta \in [0, 1]$,

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

i.e., f_i have nonnegative (upward) curvature

Source: Boyd and Vandenberghe

Convex optimization

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m \end{array}$$

Convex optimization problems are optimization problems of a special form.

- variable $x \in \mathbf{R}^n$
- equality constra
- $f_0, \ldots, f_m$ are **convex**: for $\theta \in [0, 1]$,

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

i.e., f_i have nonnegative (upward) curvature

Source: Boyd and Vandenberghe

Brief history

- theory (convex analysis): 1900–1970
- algorithms
  - 1947: simplex algorithm for linear programming (Dantzig)
  - 1960s: early interior-point methods (Fiacco & McCormick, Dikin, . . .)
  - 1970s: ellipsoid method and other subgradient methods
  - 1980s & 90s: interior-point methods (Karmarkar, Nesterov & Nemirovski)
  - since 2000s: many methods for large-scale convex optimization
- applications
  - before 1990: mostly in operations research, a few in engineering
  - since 1990: many applications in engineering (control, signal processing, communications, circuit design, . . .)
  - since 2000s: machine learning and statistics, finance

Source: Boyd and Vandenberghe

Convex

Sets

Functions

Optimization

Convex

Sets

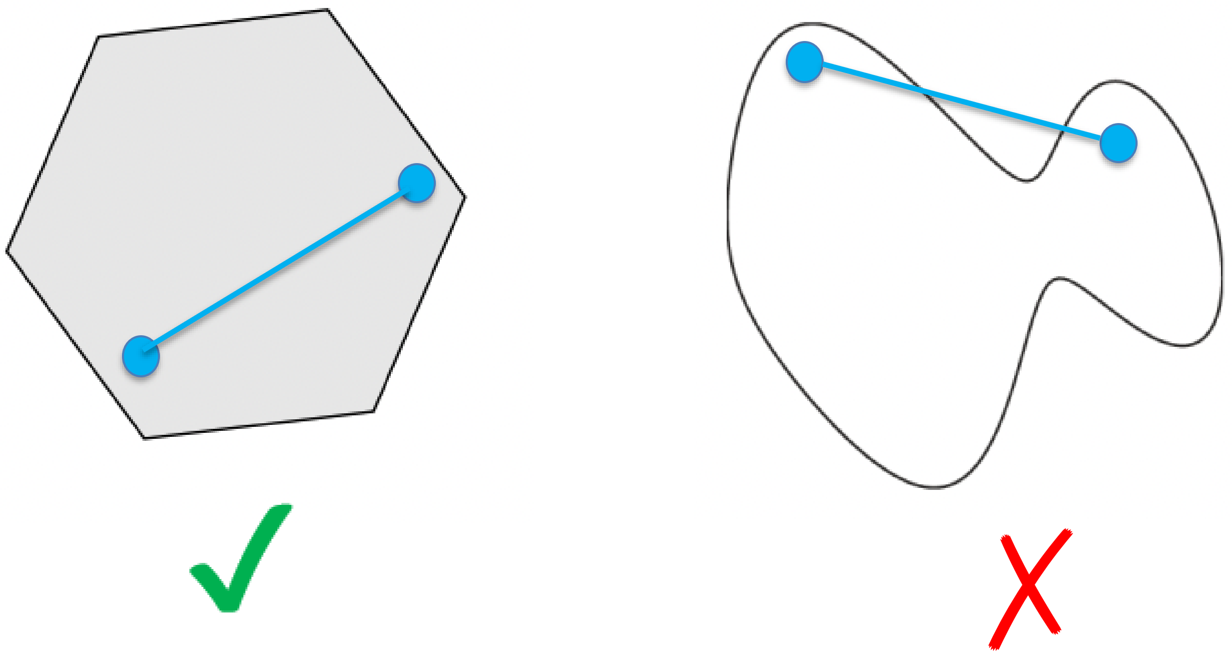
Functions

Optimization

Convex sets

convex set: contains line segment between any two points in the set

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



[Source: Boyd and Vandenberghe]

Convex combinations and convex hulls

convex combination of $x_1, \ldots, x_k$: any point x of the form

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

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

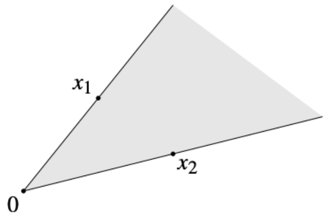
convex hull $\mathbf{conv}\, S$: set of all convex combinations of points in S

Source: Boyd and Vandenberghe

Other standard convex sets

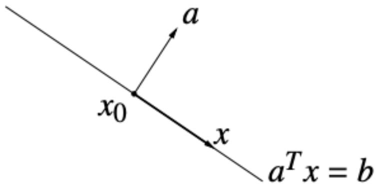
- Cones (nonnegative)

  any point of the form x = $\theta_1 x_1 + \theta_2 x_2$ with $\theta_1 \geq 0$, $\theta_2 \geq 0$



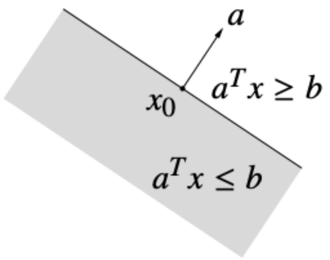
- Hyperplanes

  set of the form $\{x \mid a^T x = b\}$, with $a \neq 0$



- Halfspaces

  set of the form $\{x \mid a^T x \leq b\}$, with $a \neq 0$



Other standard convex sets

- Euclidean balls

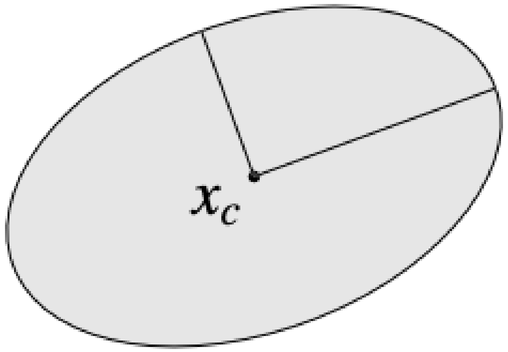
  with center x_c and radius r:

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

- Ellipsoids

  set of the form $\{x \mid (x - x_c)^T P^{-1}(x - x_c) \le 1\}$

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

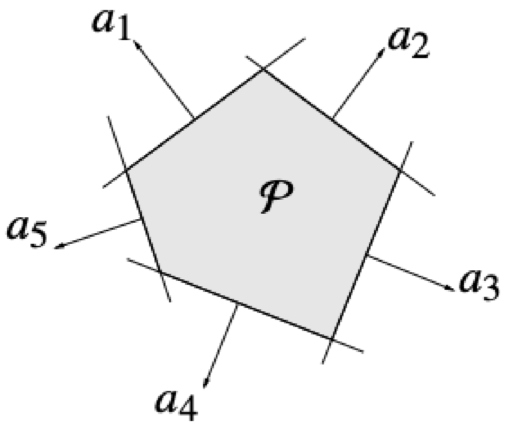


Other standard convex sets

- Polyhedra
  - solution sets of finitely many linear inequalities and equalities

    $\{x \mid Ax \preceq b, Cx = d\}$

    ($A \in \mathbf{R}^{m\times n}$, $C \in \mathbf{R}^{p\times n}$, $\preceq$ is componentwise inequality)
  - intersection of finite number of halfspaces and hyperplanes
  - example with no equality constraints; a^T_i are rows of A



Convex

Sets

Functions

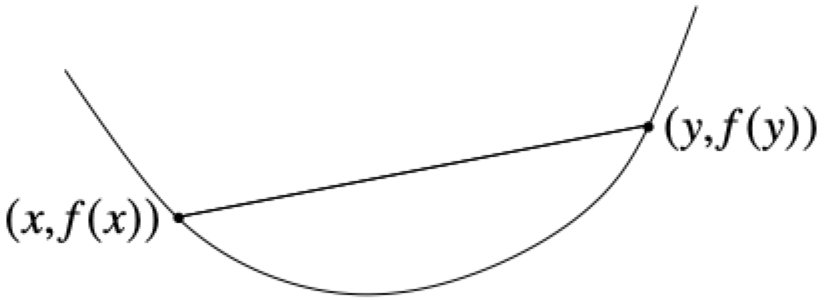
Optimization

Convex functions

Definition

- $f : \mathbf{R}^n \to \mathbf{R}$ is convex if $\mathbf{dom}\, f$ is a convex set and for all $x, y \in \mathbf{dom}\, f$, $0 \le \theta \le 1$,

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



- f is concave if $-f$ is convex
- f is strictly convex if $\mathbf{dom}\, f$ is convex and for $x, y \in \mathbf{dom}\, f$, $x \ne y$, $0 < \theta < 1$,

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

Source: Boyd and Vandenberghe

Convex functions

1-dimensional examples

- affine: $ax + b$ on $\mathbf{R}$, for any $a, b \in \mathbf{R}$
- exponential: e^{ax}, for any $a \in \mathbf{R}$
- powers: x^α on $\mathbf{R}_{++}$, for $\alpha \geq 1$ or $\alpha \leq 0$
- powers of absolute value: $|x|^p$ on $\mathbf{R}$, for $p \geq 1$
- positive part (relu): $\max\{0, x\}$

Source: Boyd and Vandenberghe

Convex functions

N-dimensional examples

- affine functions: $f(x) = a^T x + b$
- any norm, *e.g.*, the ℓ_p norms
  - $\|x\|_p = (|x_1|^p + \cdots + |x_n|^p)^{1/p}$ for $p \geq 1$
  - $\|x\|_\infty = \max\{|x_1|, \ldots, |x_n|\}$
- sum of squares: $\|x\|_2^2 = x_1^2 + \cdots + x_n^2$
- max function: $\max(x) = \max\{x_1, x_2, \ldots, x_n\}$
- softmax or log-sum-exp function: $\log(\exp x_1 + \cdots + \exp x_n)$

Source: Boyd and Vandenberghe

Convex functions

N-dimensional examples

- affine functions: $f(x) = a^T x + b$
- any norm, *e.g.*, the ℓ_p norms
  - $\|x\|_p = (|x_1|^p +$
  - $\|x\|_\infty = \max\{|x$
- sum of squares: $\|$
- max function: $\max$
- softmax or log-sum-exp function: $\log(\exp x_1 + \cdots + \exp x_n)$

Why are affine functions both convex and concave?

Creating new convex functions

- Scale by a positive constant
- Composition of two or more convex functions
- Composition with an affine function
- Pointwise maximum of two convex functions

Creating new convex functions

- Scale by a positive constant
- Composition of two or more convex functions
- Composition
- Pointwise m

Is this convex?

$f(x)=3x^2+5x+2$

Creating new convex functions

- Scale by a positive constant
- Composition of two or more convex functions
- Composition
- Pointwise m

Is this convex?

$f(x) = -e^x + 2x$

Creating new convex functions

- Scale by a positive constant
- Composition of two or more convex functions
- Composition
- Pointwise m

Is this convex?

$f(x)=e^x+x^2$

Creating new convex functions

- Scale by a positive constant
- Composition of two or more convex functions
- Composition
- Pointwise m

Is this convex?

$f(x) = 1/(x^2)$

Convex

Sets

Functions

Optimization

Principle of duality

- a mathematical property that states that two concepts or principles can be interchanged if all outcomes of one formulation are also true in the other
- solving the dual problem is sometimes easier than solving the primal problem