A Primer in Convex Optimization

Moritz Diehl
partly based on material by Colin Jones, Stephen Boyd and Lieven Vandenberghe

Overview

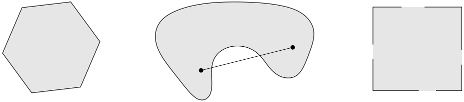
- Convex sets
- Convex functions
- Operations that preserve convexity
- Convex optimization

Convex Sets

A set $S \in \mathbb{R}^n$ is a **convex set** if for all $x_1, x_2 \in S$ and $\lambda \in [0, 1]$:

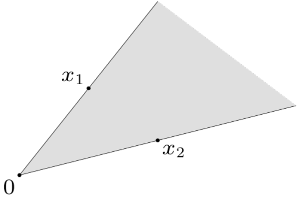
$$\lambda x_1 + (1 - \lambda) x_2 \in S$$

(set contains line segment between any two of its points)



A set $S \in \mathbb{R}^n$ is a **convex cone** if for all $x_1, x_2 \in S$ and $\theta_1, \theta_2 \geq 0$:

$$\theta_1 x_1 + \theta_2 x_2 \in S$$

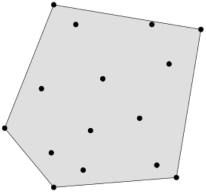


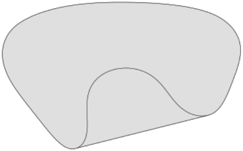
Convex hull

Convex combination of $z_1, \ldots, z_k$: Any point z of the form

$$z = \theta_1 z_1 + \theta_2 z_2 + \ldots + \theta_k z_k \text{ with } \theta_1 + \ldots + \theta_k = 1, \theta_i \geq 0$$

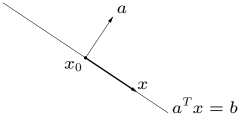
Convex hull of S: set of all convex combinations of points in S.



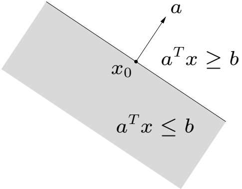


Convex sets: Hyperplanes and Halfspaces

- *Hyperplane:* Set of the form $\{x \mid a^\top x = b\}$ $(a \neq 0)$



- *Halfspace:* Set of the form $\{x \mid a^\top x \leq b\}$ $(a \neq 0)$



- Useful representation: $\left\{x \,\middle|\, a^\top (x - x_0) \leq 0\right\}$
  a is normal vector, x_0 lies on the boundary
- Hyperplanes are affine and convex, halfspaces are convex

Convex sets: Polyhedra

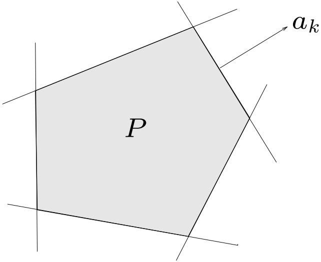
Polyhedron

A *polyhedron* is the intersection of a finite number of halfspaces.

$$P := \left\{ x \,\middle|\, a_i^\top x \le b_i,\ i = 1, \dots, n \right\}$$

A *polytope* is a bounded polyhedron.

Often written as $P := \{x \mid Ax \le b\}$, for matrix $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$, where the inequality is understood row-wise.



Operations that preserve convexity of sets

- *intersection:* the intersection of (any number of) convex sets is convex (but unification is generally non-convex)
- *affine image:* the image $f(S) := \{f(x) \mid x \in S\}$ of a convex set S under an affine function $f(x) = Ax + b$ is convex
- *affine pre-image:* the pre-image $f^{-1}(S) := \{x \mid f(x) \in S\}$ of a convex set S under an affine function $f(x) = Ax + b$ is convex

Examples

- $\left\{ x \mid x_1 + x_2 t + x_3 t^2 + x_4 t^3 \geq 0 \text{ for all } t \in [0,1] \right\}$ is convex (set of positive polynomials on unit inverval, intersection of halfspaces)
- $\{ a + Pw \mid \|w\|_2 \leq 1 \}$ is convex (affine image of unit ball)
- $\{ x \mid \|Ax + b\|_2 \leq 1 \}$ is convex (affine pre-image of unit ball)

The cone of positive semidefinite matrices

Definitions

- set of symmetric $n \times n$ matrices:
  $\mathbb{S}^n := \left\{X \in \mathbb{R}^{n\times n} \,\middle|\, X = X^\top\right\}$
- $X \succeq 0$: for all $z \in \mathbb{R}^n$ holds $z^\top X z \geq 0$ (all eigenvalues of X are non-negative)
- $X \succ 0$: all eigenvalues of X are positive
- set of positive semidefinite $n \times n$ matrices:
  $\mathbb{S}^n_+ := \{X \in \mathbb{S}^n \mid X \succeq 0\}$

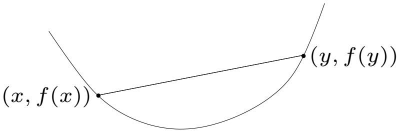
Theorem: $\mathbb{S}^n_+$ is a convex set

Proof: $\mathbb{S}^n_+ = \left\{X \in \mathbb{S}^n \,\middle|\, z^\top X z \geq 0 \text{ for all } z \in \mathbb{R}^n\right\}$ is intersection of (infinitely many) halfspaces.

Convex function: Definition

- *Convex function:*
  A function $f : S \to \mathbb{R}$ is convex if S is convex and

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y)$$
$$\text{for all } x, y \in S, \lambda \in [0,1]$$



- A function $f : S \to \mathbb{R}$ is **strictly convex** if S is convex and

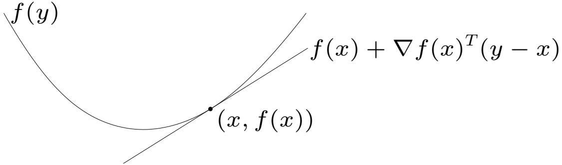
$$f(\lambda x + (1-\lambda)y) < \lambda f(x) + (1-\lambda) f(y)$$
$$\text{for all } x, y \in S, \lambda \in (0,1)$$

- A function $f : S \to \mathbb{R}$ is **concave** if $-f$ is convex.

First and second order condition for convexity

First-order condition: Differentiable f with convex domain is convex if and only if

$$f(y) \geq f(x) + \nabla f(x)^\top (y - x) \quad \text{for all } x, y \in \text{dom } f$$



Note: first-order approximation of f is global underestimator

Second-order condition: Twice differentiable f with convex domain is convex if and only if

$$\nabla^2 f(x) \succeq 0 \quad \text{for all } x \in \text{dom } f$$

Convex functions – Examples

Examples on $\mathbb{R}$:

- exponential: e^{ax}, for any $a \in \mathbb{R}$
- powers: x^a on $\mathbb{R}_+$ for $a \geq 1$ or $a \leq 0$ (otherwise concave)
- negative logarithm: $-\log x$ on $\mathbb{R}_+$

Examples on $\mathbb{R}^n$:

- affine function: $f(x) = a^\top x + b$
- norms: $\|x\|_p = (\sum_{i=1}^n |x_i|^p)^{1/p}$ for $p \geq 1$; $\|x\|_\infty = \max_k |x_k|$
- convex quadratic: $f(x) = x^\top Bx + g^\top x + c$ with $B \succeq 0$ ($\nabla^2 f(x) = 2B$)
- log-sum-exp: $f(x) = \log\left(\sum_{i=1}^n \exp(x_i)\right)$ ("smoothed max", as $\lim_{s \to 0} s\, f(x/s) = \max\{x_1, \ldots, x_n\}$)

Operations that preserve convexity of functions

- *nonnegative weighted sum:* $f(x) = \sum_{j=1}^{m} \alpha_j f_j(x)$ is convex if $\alpha_j \geq 0$ and all f_j are convex
- *composition with affine function:* $f(x) = g(Ax + b)$ is convex if g is convex
- *pointwise maximum:* $f(x) = \max\{f_1(x), \ldots, f_m(x)\}$ is convex if all f_j are convex (even supremum over infinitely many functions)
- *minimization:* if $g(x, u)$ is jointly convex in (x, u) then $f(x) = \inf_u g(x, u)$ is convex
- *convex in monotone convex:* $f(x) = h(g(x))$ is convex if g is convex and $h : \mathbb{R} \to \mathbb{R}$ is monotonely non-decreasing and convex. Proof for smooth functions: $\nabla^2 f(x) = h''(g(x))\nabla g(x)\nabla g(x)^T + h'(g(x))\nabla^2 g(x)$

Examples

- composition with affine function: $f(x) = \|Ax + b\|_2$
- expectation $f(x) = \mathbb{E}_w\{\|A(w)x + b(w)\|_2\}$ is convex (nonnegative weighted sum)
- $f(x) = \exp(c^\top x + d) - \log(a^\top x + b)$ is convex on $\{x \mid a^\top x + b > 0\}$
- pointwise maximum: $f(x) = \max_{\|w\|_2 \le 1}(a + Pw)^\top x = a^\top x + \|P^\top x\|_2$ is convex (used for robust LP)
- minimization: for $R \succ 0$, regard
$$f(x) = \min_u \begin{bmatrix} x \\ u \end{bmatrix}^\top \begin{bmatrix} Q & S^\top \\ S & R \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} = x^\top (Q - S^\top R^{-1} S)x.$$
This $f(x)$ is convex if $\begin{bmatrix} Q & S^\top \\ S & R \end{bmatrix} \succeq 0$ (cf. Schur complement)

Connecting convex sets and functions: sublevel sets

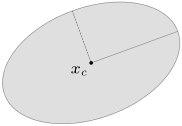
Theorem: Sublevel set $S = \{x \mid f(x) \leq c\}$ of a convex function f is a convex set

Proof: $x, y \in S$ and convexity of f imply for $t \in [0, 1]$ that $f(tx + (1 - t)y) \leq tf(x) + (1 - t)f(y) \leq c$.

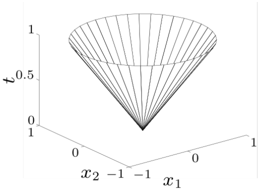
Note: the sign of the inequality matters - superlevel sets $\{x \mid f(x) \geq c\}$ would not be convex.

Convex sublevel sets – Examples

- norm balls: $\{x \in \mathbb{R}^n \mid \|x - x_c\| \leq r\}$ for any norm $\|\cdot\|$, with radius $r > 0$ and centerpoint x_c
- ellipsoids: $\left\{x \in \mathbb{R}^n \,\middle|\, (x - x_c)^\top P^{-1}(x - x_c) \leq 1\right\}$ for any positive definite shape matrix $P \succ 0$



- norm cones: $\left\{(x, t) \in \mathbb{R}^{n+1} \mid \|x\| \leq t\right\}$



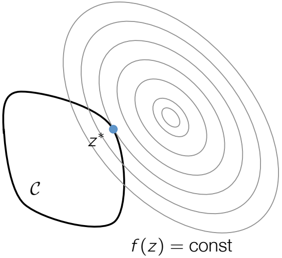
Overview

- Convex sets
- Convex functions
- Operations that preserve convexity
- **Convex optimization**

Recall: General Optimization Problem

$$\begin{aligned} \underset{z}{\text{minimize}} \quad & f(z) \\ \text{subject to} \quad & g_i(z) = 0,\ i = 1, \dots, p \\ & h_i(z) \le 0,\ i = 1, \dots, m \end{aligned}$$

- $z = (z_1, \dots, z_n)$: variables
- $f : \mathbb{R}^n \to \mathbb{R}$: objective function
- $g : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \dots, p$: equality constraint functions
- $h : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \dots, m$: inequality constraint functions



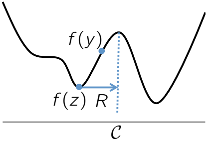
- $\mathcal{C} := \{z \mid h_i(z) \le 0,\ i = 1, \dots, m, g_i(z) = 0,\ i = 1, \dots, p\}$: *feasible set*

Optimality

minimal value: smallest possible cost $p^* := \inf\{f(z) \mid z \in \mathcal{C}\}$.
minimizer: feasible z^* with $f(z^*) = p^*$; set of all minimizers: $\{z \in \mathcal{C} \mid f(z) = p^*\}$

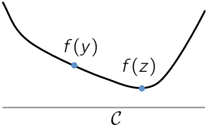
- $z \in \mathcal{C}$ is *locally optimal* if, for some $R > 0$, it satisfies

$$y \in \mathcal{C}, \|y - z\| \leq R \Rightarrow f(y) \geq f(z)$$



- $z \in \mathcal{C}$ is *globally optimal* if it satisfies

$$y \in \mathcal{C} \Rightarrow f(y) \geq f(z)$$



- If $p^* = -\infty$ the problem is *unbounded below*
- If $\mathcal{C}$ is empty, then the problem is said to be *infeasible* (convention: $p^* = \infty$)

Convex optimization problem in standard form

$$\begin{aligned} \underset{z}{\text{minimize}} \quad & f(z) \\ \text{subject to} \quad & h_i(z) \leq 0,\ i = 1, \dots, m \\ & c_i^\top z = b_i,\ i = 1, \dots, p \end{aligned}$$

- $f, h_1, \dots, h_m$ are convex
- equality constraints are affine

often rewritten as

$$\begin{aligned} \underset{z}{\text{minimize}} \quad & f(z) \\ \text{subject to} \quad & h(z) \leq 0 \\ & Cz = b \end{aligned}$$

where $C \in \mathbb{R}^{p \times n}$ and $h : \mathbb{R}^n \to \mathbb{R}^m$.

Note: With nonlinear equalities, feasible set would generally not be convex

Local and global optimality in convex optimization

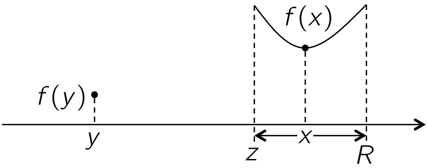
Lemma

Any locally optimal point of a convex problem is globally optimal.

Proof:

Assume x locally optimal and a feasible y such $f(y) < f(x)$.
x locally optimal implies that there exists an $R > 0$ such that

$$\|z - x\|_2 \leq R \;\Rightarrow\; f(z) \geq f(x)$$



Local and global optimality in convex optimization

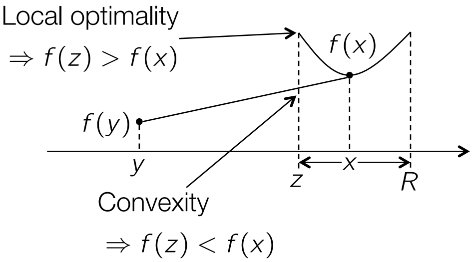
Lemma

Any locally optimal point of a convex problem is globally optimal.

Proof:

Assume x locally optimal and a feasible y such $f(y) < f(x)$.
x locally optimal implies that there exists an $R > 0$ such that

$$\|z - x\|_2 \leq R \;\Rightarrow\; f(z) \geq f(x)$$



Linear Program (LP)

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & c^\top x \\ \text{subject to} \quad & c_i^\top x + d_i \leq 0, \; i = 1, \ldots, m \\ & Ax = b \end{aligned}$$

LP Example

$$\begin{aligned} \underset{x \in \mathbb{R}^n}{\text{minimize}} \quad & \|Ax + b\|_1 \\ \text{subject to} \quad & Cx + d = 0 \end{aligned}$$

equivalent to

$$\begin{aligned} \underset{x \in \mathbb{R}^n, s \in \mathbb{R}^m}{\text{minimize}} \quad & \sum_{i=1}^{m} s_i \\ \text{subject to} \quad & -s \leq Ax + b \leq s \\ & Cx + d = 0 \end{aligned}$$

Quadratic Program (QP)

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & c^\top x + \frac{1}{2} x^\top B x \\ \text{subject to} \quad & c_i^\top x + d_i \leq 0, \ i = 1, \ldots, m \\ & Ax = b \end{aligned}$$

convex if $B \succeq 0$
strictly convex if $B \succ 0$

Quadratically Constrained Quadratic Program (QCQP)

$$
\begin{array}{ll}
\underset{x}{\text{minimize}} & x^\top B_0 x + c_0^\top x + r_0 \\
\text{subject to} & x^\top B_i x + c_i^\top x + r_i \leq 0, \; i = 1, \ldots, m \\
& Ax = b
\end{array}
$$

convex if $B_0, \ldots, B_m \succeq 0$

Second Order Cone Program (SOCP)

$$
\begin{aligned}
\underset{x}{\text{minimize}} \quad & c^\top x \\
\text{subject to} \quad & \|A_i x + b_i\|_2 \le c_i^\top x + d_i, \;\; i = 1, \ldots, m \\
& Ax = b
\end{aligned}
$$

SOCP example: robust LP

Robust LP with uncertain w:

$$\begin{array}{ll} \underset{x}{\text{minimize}} & c^\top x \\ \text{subject to} & \max\limits_{\|w\|_2 \leq 1} (a_i + P_i w)^\top x \leq b_i \;\; i = 1, \ldots, m \end{array}$$

equivalent to SOCP

$$\begin{array}{ll} \underset{x}{\text{minimize}} & c^\top x \\ \text{subject to} & a_i^\top x + \|P^\top x\|_2 \leq b_i \;\; i = 1, \ldots, m \end{array}$$

Semidefinite Program (SDP)

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & c^\top x \\ \text{subject to} \quad & x_1 F_1 + \cdots + x_n F_n + G \succeq 0 \\ & Ax = b \end{aligned}$$

with $F_1, \ldots, F_n, G \in \mathbb{S}^m$.
The generalized inequality is called **linear matrix inequality (LMI)**.

SDP Example

Eigenvalue minimization: $\underset{x \in \mathbb{R}^n}{\text{minimize}}\ \lambda_{\max}(A(x))$ with
$A(x) = A_0 + x_1 A_1 + \cdots + x_n A_n$

Equivalent SDP:

$$\begin{aligned} \underset{x \in \mathbb{R}^n, t \in \mathbb{R}}{\text{minimize}} \quad & t \\ \text{subject to} \quad & t\, I - A(x) \succeq 0 \end{aligned}$$

Proof: $t\, I \succeq A(x) \Leftrightarrow t \geq \lambda_{\max}(A(x))$

SDP comprises LP, QP, QCQP and SOCP

Among all discussed convex problem classes, SDP is most general.

Any LP can be formulated as a QP.
Any QP can be formulated as a QCQP.
Any QCQP can be formulated as a SOCP.
Any SOCP can be formulated as a SDP.

$$\text{LP} \Rightarrow \text{QP} \Rightarrow \text{QCQP} \Rightarrow \text{SOCP} \Rightarrow \text{SDP}$$

In principle, an SDP solver could be used to solve LP, QP, QCQP, SOCP and SDP... but the tailored solvers are more efficient!

Note: an NLP solver can also be used to globally solve LP, QP, or QCQP (but not for SOCP and SDP, due to non-smoothness of the generalized inequalities)

Solvers for Convex Optimization

- LP: myriads of solvers, e.g. CPLEX, GUROBI, SOPLEX
- QP: many solvers, e.g. CPLEX, OOQP, QPSOL, QPKWIK
  Embedded QP solvers: qpOASES, FORCES, HPMPC, qpDUNES, ...
- SOCP: MOSEK, ECOS
- SDP: SDPT3, sedumi

Consult "decision tree for optimization software" by Hans Mittelmann:

http://plato.la.asu.edu/guide.html

Modelling Environments for Convex Optimization

- YALMIP (from matlab)
- CVX (from matlab)
- CVXOPT (from python)
- CVXPY (from python)

Summary

- Convex optimization problem:
  - Convex cost function
  - Convex inequality constraints
  - Affine equality constraints
- main benefit of convex problems: local $=$ global optimality

Literature

- S. Boyd and L. Vandenberghe: Convex Optimization, Cambridge Univ. Press, 2004
- D. Bertsekas: Convex Optimization Theory / Convex Optimization Algorithms, Athena Scientific, 2009 / 2015