14. Convex programming

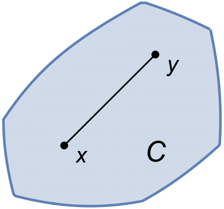
- Convex sets and functions
- Convex programs
- Hierarchy of complexity
- Example: geometric programming

Laurent Lessard (www.laurentlessard.com)

Convex sets

A set of points $C \subseteq \mathbb{R}^n$ is **convex** if for all points $x, y \in C$ and any real number $0 \leq \alpha \leq 1$, we have $\alpha x + (1 - \alpha)y \in C$.

- all points in C can **see** each other.
- can be closed or open (includes boundary or not), or some combination where only some boundary points are included.
- can be bounded or unbounded.



Convex sets

Intersections preserve convexity:
If $\mathcal{I}$ is a collection of convex sets $\{C_i\}_{i \in \mathcal{I}}$, then the intersection $S = \bigcap_{i \in \mathcal{I}} C_i$ is convex.

proof: Suppose $x, y \in S$ and $0 \le \alpha \le 1$. By definition, $x, y \in C_i$ for each $i \in \mathcal{I}$. By the convexity of C_i, we must have $\alpha x + (1 - \alpha) y \in C_i$ as well. Therefore $\alpha x + (1 - \alpha) y \in S$, and we are done.

note: The union of convex sets $C_1 \cup C_2$ is need not be convex!

Convex sets

Constraints can be characterized by sets!

- If we define $C_1 := \{x \in \mathbb{R}^n \mid Ax \leq b\}$ then:

$$Ax \leq b \quad \Longleftrightarrow \quad x \in C_1$$

- If we define $C_2 := \{x \in \mathbb{R}^n \mid Fx = g\}$ then:

$$Ax \leq b \text{ and } Fx = g \quad \Longleftrightarrow \quad x \in C_1 \cap C_2$$

Convex sets

Example: SOCP

Let $C := \left\{x \in \mathbb{R}^n \;\middle|\; \|Ax + b\| \le c^\mathsf{T} x + d\right\}$. To prove C is convex, suppose $x, y \in C$ and let $z := \alpha x + (1-\alpha)y$. Then:

$$\begin{aligned}
\|Az + b\| &= \|A(\alpha x + (1-\alpha)y) + b\| \\
&= \|\alpha(Ax + b) + (1-\alpha)(Ay + b)\| \\
&\le \alpha\|Ax + b\| + (1-\alpha)\|Ay + b\| \\
&\le \alpha(c^\mathsf{T} x + d) + (1-\alpha)(c^\mathsf{T} y + d) \\
&= c^\mathsf{T} z + d
\end{aligned}$$

Therefore, $\|Az + b\| \le c^\mathsf{T} z + d$, i.e. C is convex.

Convex sets

Example: spectrahedron

Let $C := \left\{ x \in \mathbb{R}^3 \;\middle|\; \begin{bmatrix} 1 & x_1 & x_2 \\ x_1 & 1 & x_3 \\ x_2 & x_3 & 1 \end{bmatrix} \succeq 0 \right\}$. To prove C is convex, consider the set $S := \left\{ X \in \mathbb{R}^{3\times 3} \;\middle|\; X = X^\mathsf{T} \succeq 0 \right\}$ Note that S is the PSD cone. It is convex because if we define $Z := \alpha X + (1-\alpha)Y$ where $X, Y \in S$ and $0 \le \alpha \le 1$, then

$$\begin{aligned} w^\mathsf{T} Z w &= w^\mathsf{T} \left(\alpha X + (1-\alpha) Y \right) w \\ &= \alpha w^\mathsf{T} X w + (1-\alpha) w^\mathsf{T} Y w \end{aligned}$$

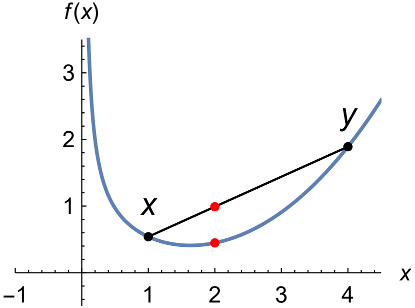
So if $X \succeq 0$ and $Y \succeq 0$, then $Z \succeq 0$. So S is convex. Now, C is convex because it's the intersection of two convex sets: the PSD cone S and the affine space $\{X \in \mathbb{R}^{3\times 3} \mid X_{ii} = 1\}$.

Convex functions

- If $C \subseteq \mathbb{R}^n$, a function $f : C \to \mathbb{R}$ is **convex** if C is a convex set and for all $x, y \in C$ and $0 \leq \alpha \leq 1$, we have:

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y)$$

- f is **concave** if $-f$ is convex.



Convex and concave functions

Convex functions on $\mathbb{R}$:

- Affine: $ax + b$.
- Absolute value: $|x|$.
- Quadratic: ax^2 for any $a \geq 0$.
- Exponential: a^x for any $a > 0$.
- Powers: x^α for $x > 0$, $\alpha \geq 1$ or $\alpha \leq 0$.
- Negative entropy: $x \log x$ for $x > 0$.

Concave functions on $\mathbb{R}$:

- Affine: $ax + b$.
- Quadratic: ax^2 for any $a \leq 0$.
- Powers: x^α for $x > 0$, $0 \leq \alpha \leq 1$.
- Logarithm: $\log x$ for $x > 0$.

Convex and concave functions

Convex functions on $\mathbb{R}^n$:

- Affine: $a^\mathsf{T} x + b$.
- Norms: $\|x\|_2$, $\|x\|_1$, $\|x\|_\infty$
- Quadratic form: $x^\mathsf{T} Q x$ for any $Q \succeq 0$

Building convex functions

1. Nonnegative weighted sum: If $f(x)$ and $g(x)$ are convex and $\alpha, \beta \geq 0$, then $\alpha f(x) + \beta g(x)$ is convex.

2. Composition with an affine function:
If $f(x)$ is convex, so is $g(x) := f(Ax + b)$

3. Pointwise maximum: If $f_1(x), \ldots, f_k(x)$ are convex, then $g(x) := \max\{f_1(x), \ldots, f_k(x)\}$ is convex.

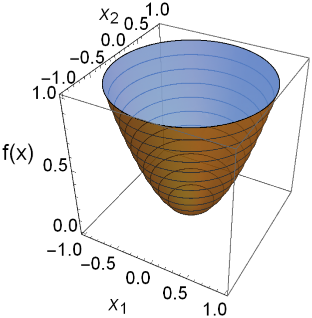
proof: Let $z := \alpha x + (1 - \alpha)y$ as usual.

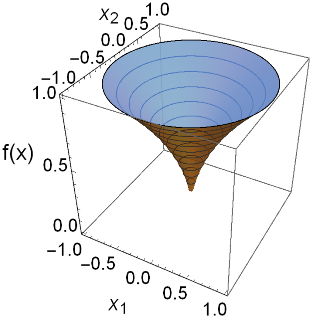
$$\begin{aligned} g(z) &= f(Az + b) \\ &= f(\alpha(Ax + b) + (1 - \alpha)(Ay + b)) \\ &\leq \alpha f(Ax + b) + (1 - \alpha)f(Ay + b) \\ &= \alpha g(x) + (1 - \alpha)g(y) \end{aligned}$$

Convex functions vs sets

Level set: If f is a convex function, then the set of points satisfying $f(x) \leq a$ is a convex set.

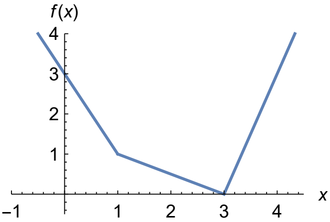
- Converse is false: if all level sets of f are convex, it does not necessarily imply that f is a convex function!



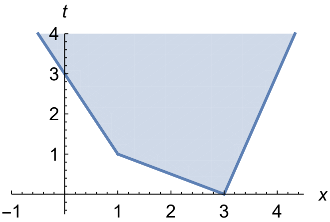


Convex functions vs sets

Epigraph: $f : \mathbb{R}^n \to \mathbb{R}$ is a convex function **if and only if** the set $\{(x, t) \in \mathbb{R}^{n+1} \mid f(x) \leq t\}$ is convex.



plot of $f(x)$



$\{(x, t) \mid f(x) \leq t\}$

Convex programs

The standard form for a convex optimization problem:

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & f_0(x) \\ \text{subject to:} \quad & f_i(x) \le 0 \quad \text{for } i = 1, \dots, k \\ & Ax = b \\ & x \in C \end{aligned}$$

- $f_0, f_1, \dots, f_k$ are convex functions
- C is a convex set

Convex programs

- Can turn $f_0(x)$ into a linear constraint (use epigraph)
- Can characterize constraints using sets.

Minimalist form:

$$\underset{x \in S}{\text{minimize}} \quad c^\mathsf{T} x$$

- S is a convex set

Key properties and advantages

1. The set of optimal points $x^\star$ is itself a convex set.
   - **Proof**: If $x^\star$ and $y^\star$ are optimal, then we must have $f^\star = f_0(x^\star) = f_0(y^\star)$. Also, $f^\star \leq f_0(z)$ for any z. Choose $z := \alpha x^\star + (1-\alpha)y^\star$ with $0 \leq \alpha \leq 1$. By convexity of f_0, $f^\star \leq f_0(\alpha x^\star + (1-\alpha)y^\star) \leq \alpha f_0(x^\star) + (1-\alpha)f_0(y^\star) = f^\star$. Therefore, $f_0(z) = f^\star$, i.e. z is also an optimal point.
2. If x is a locally optimal point, then it is globally optimal.
   - Follows from the result above. A very powerful fact!
3. Upper an lower bounds available via duality (more later!)
4. Often numerically tractable (not always!)

Hierarchy of programs

From least general to most general model:

1. LP: linear cost and linear constraints
2. QP: convex quadratic cost and linear constraints
3. QCQP: convex quadratic cost and constraints
4. SOCP: linear cost, second order cone constraints
5. SDP: linear cost, semidefinite constraints
6. CVX: convex cost and constraints

Less general (simpler) models are typically preferable

Solving convex problems

Simpler models are usually more efficient to solve

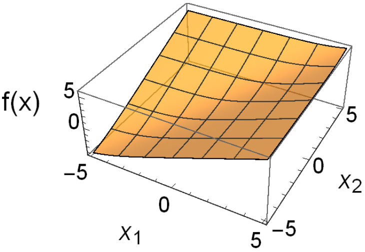
Factors affecting solver speed:

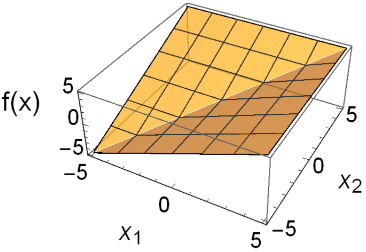
- How difficult is it to verify that $x \in C$?
- How difficult is it to project onto C ?
- How difficult is it to evaluate $f(x)$?
- How difficult is it to compute $\nabla f(x)$?
- Can the solver take advantage of sparsity?

Example: geometric programming

The **log-sum-exp** function (shown left) is convex:

$$f(x) := \log \left(\sum_{k=1}^{n} \exp x_k \right)$$





It's a smoothed version of $\max\{x_1, \ldots, x_k\}$ (shown right)

Example: geometric programming

Suppose we have positive decision variables $x_i > 0$, and constraints of the form (with each $c_j > 0$ and $\alpha_{jk} \in \mathbb{R}$):

$$\sum_{j=1} c_j x_1^{\alpha_{j1}} x_2^{\alpha_{j2}} \cdots x_n^{\alpha_{jn}} \leq 1$$

Then by using the substitution $y_i := \log(x_i)$, we have:

$$\log\left(\sum_{j=1}^{n} \exp\left(a_{j0} + a_{j1}y_1 + \cdots + a_{jn}y_n\right)\right) \leq 0$$

(where $a_{j0} := \log c_j$). This is a log-sum-exp function composed with an affine function (convex!)

Example: geometric programming

Example: We want to design a box of height h, width w, and depth d with maximum volume (hwd) subject to the limits:

- total wall area: $2(hw + hd) \leq A_{\text{wall}}$
- total floor area: $wd \leq A_{\text{flr}}$
- height-width aspect ratio: $\alpha \leq \frac{h}{w} \leq \beta$
- width-depth aspect ratio: $\gamma \leq \frac{d}{w} \leq \delta$

We can make some of the constraints linear, but not all of them. This appears to be a nonconvex optimization problem...

Example: geometric programming

Example: We want to design a box of height h, width w, and depth d with maximum volume (hwd) subject to the limits:

- total wall area: $2(hw + hd) \leq A_{\text{wall}}$
- total floor area: $wd \leq A_{\text{flr}}$
- height-width aspect ratio: $\alpha \leq \frac{h}{w} \leq \beta$
- width-depth aspect ratio: $\gamma \leq \frac{d}{w} \leq \delta$

$$
\begin{array}{lll}
\underset{h,w,d>0}{\text{minimize}} & h^{-1}w^{-1}d^{-1} & \\
\text{subject to:} & \frac{2}{A_{\text{wall}}}hw + \frac{2}{A_{\text{wall}}}hd \leq 1, & \frac{1}{A_{\text{flr}}}wd \leq 1 \\
& \alpha h^{-1}w \leq 1, & \frac{1}{\beta}hw^{-1} \leq 1 \\
& \gamma wd^{-1} \leq 1, & \frac{1}{\delta}w^{-1}d \leq 1
\end{array}
$$

Example: geometric programming

$$\begin{array}{lll} \underset{h,w,d>0}{\text{minimize}} & h^{-1}w^{-1}d^{-1} & \\ \text{subject to:} & \frac{2}{A_{\text{wall}}}hw + \frac{2}{A_{\text{wall}}}hd \le 1, & \frac{1}{A_{\text{flr}}}wd \le 1 \\ & \alpha h^{-1}w \le 1, & \frac{1}{\beta}hw^{-1} \le 1 \\ & \gamma wd^{-1} \le 1, & \frac{1}{\delta}w^{-1}d \le 1 \end{array}$$

- Define: $x := \log h$, $y := \log w$, and $z := \log d$.
- Express the problem in terms of the new variables x, y, z. Note: h, w, d are positive but x, y, z are unconstrained.

Example: geometric programming

$$
\begin{aligned}
\underset{x,y,z}{\text{minimize}} \quad & \log(e^{-x-y-z}) \\
\text{subject to:} \quad & \log(e^{\log(2/A_{\text{wall}})+x+y} + e^{\log(2/A_{\text{wall}})+x+z}) \le 0 \\
& \log(e^{\log(1/A_{\text{flr}})+y+z}) \le 0 \\
& \log(e^{\log\alpha - x+y}) \le 0, \qquad \log(e^{-\log\beta + x-y}) \le 0 \\
& \log(e^{\log\gamma + y-z}) \le 0, \qquad \log(e^{-\log\delta - y+z}) \le 0
\end{aligned}
$$

- this is a convex model, but it can be simplified!
- most of the constraints are actually linear.

Example: geometric programming

$$
\begin{aligned}
\underset{x,y,z}{\text{minimize}} \quad & -x-y-z \\
\text{subject to:} \quad & \log\left(e^{\log(2/A_{\text{wall}})+x+y} + e^{\log(2/A_{\text{wall}})+x+z}\right) \le 0 \\
& y+z \le \log A_{\text{flr}} \\
& \log\alpha \le x-y \le \log\beta \\
& \log\gamma \le z-y \le \log\delta
\end{aligned}
$$

- This is a convex optimization problem.

For the curious...

- We saw several types of cones, each of which corresponds to a convex program:
  - polyhedral cone: LP
  - quadratic cone: QP and QCQP
  - second order cone: SOCP
  - semidefinite cone: SDP
- Geometric programs are actually also cone programs! They can be expressed using the **exponential cone**, which is

$$\left\{(x, y, z) \,\middle|\, z/y \geq e^{x/y} \text{ and } y, z > 0\right\}$$

- Exponential cones are not comparable to semidefinite cones (neither a subset nor a superset).

For the curious...

- There are many types of cones. A good overview:
  https://docs.mosek.com/modeling-cookbook/index.html

- Every convex problem we've seen so far in this class can be written as a **convex cone program**

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & c^\mathsf{T} x \\ \text{subject to:} \quad & Ax = b \\ & x \in \mathcal{K} \end{aligned}$$

  where $\mathcal{K}$ is a Cartesian product of cones.

- Solvers that handle popular cones: SCS, COSMO, Mosek

For the curious...

To solve a general convex problem that can't be expressed as a cone program (rare), use the Ipopt solver.

- Ipopt can also solve nonconvex optimization problems.
- It will return termination status `LOCALLY_SOLVED` when successful. In general, the solution it finds depends on how the solver is initialized (more on this later when we look at nonconvex problems).
- If you give Ipopt a convex optimization problem, it will still say `LOCALLY_SOLVED`, but since the problem is convex, you know this is actually a global solution!