13. Convex programming

- Convex sets and functions
- Convex programs
- Hierarchy of complexity
- Example: geometric programming
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 $I$ is a collection of convex sets $\{C_i\}_{i \in I}$, then the intersection $S = \bigcap_{i \in I} C_i$ is convex.

**proof:** Suppose $x, y \in S$ and $0 \leq \alpha \leq 1$. By definition, $x, y \in C_i$ for each $i \in I$. 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:
  \[ Ax \leq b \iff x \in C_1 \]

- If we define $C_2 := \{ x \in \mathbb{R}^n \mid Fx = g \}$ then:
  \[ Ax \leq b \text{ and } Fx = g \iff x \in C_1 \cap C_2 \]
Convex sets

**Example: SOCP**

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

\[
\|Az + b\| = \|A(\alpha x + (1 - \alpha)y) + b\| \\
= \|\alpha(Ax + b) + (1 - \alpha)(Ay + b)\| \\
\leq \alpha\|Ax + b\| + (1 - \alpha)\|Ay + b\| \\
\leq \alpha(c^T x + d) + (1 - \alpha)(c^T y + d) \\
= c^T z + d
\]

Therefore, $\|Az + b\| \leq c^T z + d$, i.e. $C$ is convex.
Convex sets

Example: spectrahedron

Let \( C := \left\{ x \in \mathbb{R}^3 \mid \begin{bmatrix} 1 & x_1 & x_2 \\ x_1 & 1 & x_3 \\ x_2 & x_3 & 1 \end{bmatrix} \succeq 0 \right\} \). To prove \( C \) is convex, consider the set \( S := \left\{ X \in \mathbb{R}^{3 \times 3} \mid X = X^T \succeq 0 \right\} \)

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 \( 0 \leq \alpha \leq 1 \), then

\[
 w^T Z w = w^T (\alpha X + (1 - \alpha) Y) w = \alpha w^T X w + (1 - \alpha) w^T Y w
\]

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: $x^\alpha$ 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: $x^\alpha$ for $x > 0$, $0 \leq \alpha \leq 1$.
- Logarithm: $\log x$ for $x > 0$. 
Convex functions on $\mathbb{R}^n$:

- **Affine:** $a^T x + b$.

- **Norms:** $\|x\|_2$, $\|x\|_1$, $\|x\|_\infty$

- **Quadratic form:** $x^T Q x$ 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.

\[
g(z) = f(Az + b) \\
= f(\alpha(Ax + b) + (1 - \alpha)(Ay + b)) \\
\leq \alpha f(Ax + b) + (1 - \alpha)f(Ay + b) \\
= \alpha g(x) + (1 - \alpha)g(y)
\]
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.

- If all level sets are convex, then $f$ is not necessarily convex!
Convex functions vs sets

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

The standard form for a convex optimization problem:

\[
\begin{align*}
\text{minimize} \quad & f_0(x) \\
\text{subject to:} \quad & f_i(x) \leq 0 \quad \text{for } i = 1, \ldots, k \\
& Ax = b \\
& x \in C
\end{align*}
\]

- \(f_0, f_1, \ldots, f_k\) 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:

\[
\begin{align*}
\text{minimize} & \quad c^T x \\
\text{subject to} & \quad x \in S
\end{align*}
\]

- $S$ is a convex set
Key properties and advantages

1. The set of optimal points $x^*$ is itself a convex set.
   
   ▶ **Proof**: If $x^*$ and $y^*$ are optimal, then we must have $f^* = f_0(x^*) = f_0(y^*)$. Also, $f^* \leq f_0(z)$ for any $z$. Choose $z := \alpha x^* + (1 - \alpha)y^*$ with $0 \leq \alpha \leq 1$. By convexity of $f_0$, $f^* \leq f_0(\alpha x^* + (1 - \alpha)y^*) \leq \alpha f_0(x^*) + (1 - \alpha)f_0(y^*) = f^*$. Therefore, $f_0(z) = f^*$, 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 and 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!
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)
Suppose we have positive decision variables $x_i > 0$, and constraints of the form (with each $c_j > 0$ and $\alpha_{jk} \in \mathbb{R}$):

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

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: $\gamma \leq \frac{d}{w} \leq \delta$

\[
\begin{align*}
\text{minimize } & h^{-1}w^{-1}d^{-1} \\
\text{subject to: } & \frac{2}{A_{\text{wall}}} hw + \frac{2}{A_{\text{wall}}} hd \leq 1, \quad \frac{1}{A_{\text{flr}}} wd \leq 1 \\
& \alpha h^{-1}w \leq 1, \quad \frac{1}{\beta} hw^{-1} \leq 1 \\
& \gamma wd^{-1} \leq 1, \quad \frac{1}{\delta} w^{-1}d \leq 1
\end{align*}
\]
Example: geometric programming

\[
\begin{align*}
\text{minimize} & \quad h^{-1}w^{-1}d^{-1} \\
\text{subject to:} & \quad \frac{2}{A_{\text{wall}}} hw + \frac{2}{A_{\text{wall}}} hd \leq 1, \quad \frac{1}{A_{\text{flr}}} wd \leq 1 \\
& \quad \alpha h^{-1} w \leq 1, \quad \frac{1}{\beta} hw^{-1} \leq 1 \\
& \quad \gamma wd^{-1} \leq 1, \quad \frac{1}{\delta} w^{-1} d \leq 1
\end{align*}
\]

- 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

\[
\begin{align*}
\text{minimize } \quad & \log(e^{-x-y-z}) \\
\text{subject to: } \quad & \log(e^{\log(2/A_{\text{wall}}) + x + y} + e^{\log(2/A_{\text{wall}}) + x + z}) \leq 0 \\
& \log(e^{\log(1/A_{\text{flr}}) + y + z}) \leq 0 \\
& \log(e^{\log \alpha - x + y}) \leq 0, \quad \log(e^{-\log \beta + x - y}) \leq 0 \\
& \log(e^{\log \gamma + y - z}) \leq 0, \quad \log(e^{-\log \delta - y + z}) \leq 0
\end{align*}
\]

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

$$\begin{align*}
\text{minimize} & \quad -x - y - z \\
\text{subject to:} & \quad \log(e^{\log(2/A_{\text{wall}}) + x + y} + e^{\log(2/A_{\text{wall}}) + x + z}) \leq 0 \\
& \quad y + z \leq \log A_{\text{flr}} \\
& \quad \log \alpha \leq x - y \leq \log \beta \\
& \quad \log \gamma \leq z - y \leq \log \delta
\end{align*}$$

- This is a convex optimization problem.