Topic 3  Convex optimization problems

1. A general optimization problem:

\[
\begin{align*}
\min & \quad f_0(x) \\
\text{s.t.} & \quad f_i(x) \leq 0 \quad i = 1, \ldots, m \\
& \quad h_i(x) = 0 \quad i = 1, \ldots, p
\end{align*}
\]

Terminology:

- \( x \in \mathbb{R}^n \) : optimization variable
- \( f_0 : \mathbb{R}^n \rightarrow \mathbb{R} \) : objective function (or cost function)
- \( f_i : \mathbb{R}^n \rightarrow \mathbb{R} \) : inequality constraint functions
- \( h_i : \mathbb{R}^n \rightarrow \mathbb{R} \) : equality constraint functions

If there are no constraints, we say the problem is unconstrained.

\[ D = \bigcap_{i=1}^m \text{dom } f_i \cap \bigcap_{i=1}^p \text{dom } h_i \] is the domain of the optimization problem.

\( x \in D \) is a feasible point if it satisfies all the constraints.

If there exists at least one feasible point, the optimization problem is said to be feasible.

Otherwise it is infeasible (no solution exists).

Set of all feasible points is called the feasible set or the constraint set.

Optimal value \( p^* \) of the problem is the optimal objective function:

\[
p^* = \min \{ f_0(x) | f_i(x) \leq 0, i = 1, \ldots, m, h_i(x) = 0, i = 1, \ldots, p \}.
\] if the problem is infeasible \( \Rightarrow p^* = \infty \) then problem is unbounded.
Optimal point $x^*$ is a feasible point such that
\[ f_0(x^*) = p^* \]
The set of all optimal points is called the optimal set $X_{opt}$.
- If there exists at least one optimal point, the problem is solvable and the optimal value is attained or achieved.
- If $X_{opt}$ is empty then the problem is not attained or achieved (for example, when the problem is unbounded).

$x$ is locally optimal if it minimizes $f_0$ over nearby points in the feasible set.
That is, $\exists R > 0$ such that $x$ is optimal for problem
\[
\begin{align*}
\min & \quad f_2(x) \\
\text{s.t.} & \quad f_i(x) \leq 0, \quad i = 1, \ldots, m \\
& \quad h_i(x) = 0, \quad i = 1, \ldots, p \\
& \quad \|x - x^0\|_2 \leq R
\end{align*}
\]

"Globally optimal" is when $x$ solves the original optimization problem. In this course, we use optimal to mean globally optimal.

Examples:
1) $f_0(x) = \frac{1}{x} \quad p^* = 0$ but $x^*$ is not achieved.
2) $f_0(x) = -\log x \quad p^* = -\infty$, problem unbounded below.
3) $f_0(x) = x\log x \quad p^* = -\frac{1}{e} \Rightarrow x^* = \frac{1}{e}$

All above 3 examples have domain $f_0 = R^+ \cup \{0\}$, $x > 0$.

4) $f_0(x) = x^3 - 3x$ has a local optimal point $x^* = 1$.

Implicit constraints: Note that even when an optimization problem is unconstrained, it may still have implicit constraints as a result of its domain.
Example: \( \min \ f(x) = -\frac{1}{2} \log(\log x + b_i) \)

is an unconstrained problem with implicit constraints as:

\[ \frac{1}{x} - b_i \leq 0 \]
Example: \( \min f_0(x) = -\sum_{i=1}^{k} \log (a_i^T x + b_i) \) is an unconstrained problem with implicit constraints as \( a_i^T x - b_i \leq 0 \).

Feasibility problem: This kind of problems is to determine if a set of constraints is feasible (or consistent).

A feasibility problem can be written as

\[
\begin{align*}
\text{find } & \quad x \\
\text{s.t.} & \quad f_i(x) \leq 0, \quad i = 1, \ldots, m \quad \text{OR} \quad f_i(x) = 0, \quad i = 1, \ldots, p
\end{align*}
\]

2. Convex optimization problems.

\[
\begin{align*}
\min & \quad f_0(x) \\
\text{s.t.} & \quad f_i(x) \leq 0, \quad i = 1, \ldots, m \\
& \quad a_i^T x = b_i, \quad i = 1, \ldots, p
\end{align*}
\]

where \( f_0, f_1, \ldots, f_m \) are convex function. The equality constraints are affine.

If \( f_0(x) \) is quasiconvex, the problem is quasiconvex (other \( f_i(x) \) are still convex).

The feasible set of a convex optimization problem is convex,

\[
D = \bigcap_{i=0}^{m} \text{dom } f_i \cap \bigcap_{i=1}^{p} \text{dom } h_i
\]

Sublevel sets \( \{ x | f_i(x) \leq 0 \} \) and hyperplanes \( \{ x | a_i^T x = b \} \).
Note that this definition of convex problem is strict. A more general definition, or abstract convex problem, is that minimizes a convex function over a convex set. As it turns out, solving such an abstract convex optimization problem requires us to find a description of the set in terms of convex inequalities and linear equality constraints. For this reason, we focus on this restricted form of convex problems.

Example: \[\min \ x_1 + x_2\]
\[\text{s.t.} \ -x_1 \leq 0\]
\[-x_2 \leq 0\]
\[1 - x_1 x_2 \leq 0.\]

This problem is not convex in the standard form since \(f(x) = 1 - x_1 x_2\) is not convex on \(\mathbb{R}^2_+\).

But this problem can be cast as a standard convex problem:
\[\min \ x_1 + x_2\]
\[\text{s.t.} \ -x_1 \leq 0\]
\[-x_2 \leq 0\]
\[1 - \sqrt{x_1 x_2} \leq 1 \ (\text{convex})\]

\[\mathbf{x}^* = (1, 1)\]
\[\mathbf{p}^* = 2.\]

+ A convex optimization problem can also be written as \[\min_{x \in \mathcal{X}} \ f_0(x)\]
\[\text{s.t.} \ f_i(x) \leq 0, \quad i = 1, \ldots, m\]
\[Ax = b\]
\[\mathbf{A} \in \mathbb{R}^{m \times n}, \ b \in \mathbb{R}^m.\]
+ Local and global optimality:

For a convex optimization problem, a fundamental property is that any locally optimal point is also globally optimal.

This property is easy to see:

By contradiction, suppose there exists a point $y$ outside the local circle of radius $R$ centered at $x$, such that $f(y) < f_0(x)$. Take a small step from $x$ towards $y$: $z = (1-\theta)x + \theta y$ such that $z$ is still inside the circle. Then $f_0(z) \leq (1-\theta)f_0(x) + \theta f_0(y) < f_0(x)$ — contradictory to the assumption that $x$ is locally optimal.

+ Optimality criteria for differentiable $f$:

Recall the first-order condition of convex functions:

$$f(y) \geq f(x) + \nabla f(x)^T (y-x) \quad \forall x,y \in \text{dom} f.$$

Then for $x^*$ to be optimal, it must satisfy

$$\nabla f(x^*)^T (y-x^*) \geq 0 \quad \forall y \text{ feasible}.$$

The converse is also true (i.e. this is an iff statement).

This condition means that:

either $\nabla f(x^*) = 0$

or $-\nabla f(x^*)$ defines a supporting hyperplane to the feasible set at $x^*$. 

$$z = (1-\theta)x + \theta y$$

$$\theta = \frac{R}{2 \|y-x\|_2}$$
Special cases:

- **Unconstrained problems**: The optimality conditions reduce to the well-known condition

  \[ x^* \text{ optimal } \iff \nabla f(x^*) = 0. \]

  This is because for unconstrained problems, the feasible set is the domain of \( f_0(x) \).

  For differentiable \( f_0(x) \) then its domain is an open set. Therefore all points surrounding \( x^* \) and sufficiently close are feasible.

  Thus let \( y = x^* + t \nabla f_0(x^*) \), for \( t > 0 \) and small enough then \( y \in \text{dom} f_0 \):

  \[ 0 \leq \nabla f_0(x)^T (y - x^*) = -t \| \nabla f_0(x^*) \|_2^2 \]

  \[ \Rightarrow \nabla f_0(x^*) = 0. \]

**Example**: \( \min f_0(x) = -\sum_{i=1}^m \log (b_i - a_i^T x), \ a_i \in \mathbb{R}^n \).

\( \text{dom} f_0 = \{ x \mid Ax < b \} \) is an open set.

Thus the optimality conditions are:

\[ Ax < b \quad \Rightarrow \quad \nabla f_0(x) = \sum_{i=1}^m \frac{1}{b_i - a_i^T x} a_i = 0 \]

Then we solve this set of equations and inequalities to find \( x^* \).

- **Problems with equality constraints only**:

  \[ \min f_0(x) \]

  s.t. \( Ax = b \), \( A \in \mathbb{R}^{p \times n} \)

  The optimality condition is

  \[ \nabla f_0(x^*)^T (y - x^*) \geq 0 \quad \text{for all } y = A x + w \quad \text{for } w \in \mathbb{N}(A), \]

  since \( x, y \) are feasible \( \Rightarrow y = x + w \) for \( w \in \mathbb{N}(A) \), the null space of \( A \).

  Then the optimality condition can be written as
\[ \nabla f_0(x)^T \omega \geq 0 \quad \forall \omega \in N(A) \]
\[ \Rightarrow \nabla f_0(x)^T \omega = 0 \quad \forall \omega \in N(A) \quad \text{(since } N(A) \text{ is a subspace, which contains } \omega) \]

Thus \( \nabla f_0(x) \perp N(A) \) or \( \nabla f_0(x) \in N(A)^\perp = R(A^T) \).

That is, \( \nabla f_0(x) \) can be written as
\[ \nabla f_0(x) = A^T \nu \text{ for some } \nu \in \mathbb{R}^p. \]

In other words, \( x \) is optimal iff there exists a \( \nu \in \mathbb{R}^p \) such that
\[ \nabla f_0(x) + A^T \nu = 0 \]
\[ A \times x = b. \]

This is the classical Lagrange multiplier (more in the duality topic).

+ Equivalent convex problems:

Two optimization problems are equivalent if from a solution of one, a solution of the other is readily found, and vice versa.

Next we describe some general transformations that yield equivalent problems.

The first two are general transformations. The others are transformations that also preserve convexity.

(i) Change of variable:

Suppose \( \Phi : \mathbb{R}^n \rightarrow \mathbb{R}^n \) is a one-to-one function with image covering the problem domain \( D \), let \( x = \Phi(z) \) and define
\[ f_i(z) = f_i(\Phi(z)) \quad i = 0, \ldots, m \]
\[ h_i(z) = h_i(\Phi(z)) \quad i = 1, \ldots, p \]
\[ \nabla f_0(x)^T w \geq 0 \quad \forall \ w \in N(A) \]
\[ \Rightarrow \ \nabla f_0(x)^T w = 0 \quad \forall \ w \in N(A) \] (since \( N(A) \) is a subspace, which contains \( w \).
Thus
\[ \nabla f_0(x) \perp N(A) \] or \( \nabla f_0(x) \in N(A)^\perp = R(A^T).
That is, \( \nabla f_0(x) \) can be written as
\[ \nabla f_0(x) = A^T u \] for some \( u \in \mathbb{R}^p \).
In other words, \( x \) is optimal iff there exists a \( u \in \mathbb{R}^p \) such that
\[ \nabla f_0(x) + A^T u = 0 \]
\[ A x = b \]
This is the classical Lagrange multiplier (more in the duality topic).

Lecture 9:
1) Equivalent convex problems:

Two optimization problems are equivalent if from a solution of one, a solution of the other is readily found, and vice versa.

Next we describe some general transformation that yield equivalent problems.
The first two are general transformations. The others are transformation that also preserve convexity.

(i) **Change of variable**
Suppose \( \phi : \mathbb{R}^n \rightarrow \mathbb{R}^n \) is a one-to-one function with image covering the problem domain \( D \). Let \( x = \phi(z) \) and define
\[ f_i(z) = f_i(\phi(z)) \quad i = 0, \ldots, m \]
\[ h_i(z) = h_i(\phi(z)) \quad i = 1, \ldots, p \]
then the following problem is equivalent to the original problem:

\[
\begin{align*}
\min & \quad f_0(x) \\
\text{s.t.} & \quad f_i(x) \leq 0 \quad i = 1, \ldots, m \\
& \quad h_i(x) = 0 \quad i = 1, \ldots, p
\end{align*}
\]

(iii) Transformation of objective and constraint functions:

Suppose that

\[\Phi_0 : \mathbb{R} \to \mathbb{R}\] is monotone increasing
\[\Phi_i \ldots \Phi_m : \mathbb{R} \to \mathbb{R}\] satisfy \[\Phi_i(u) \leq 0 \iff u \leq 0\]
\[\Phi_{m+1} \ldots \Phi_{mp} : \mathbb{R} \to \mathbb{R}\] satisfy \[\Phi_i(u) = 0 \iff u = 0\]

then we can define

\[\tilde{f}_i(x) = \Phi_i(f_i(x)) \quad i = 0, \ldots, m\]
\[\tilde{h}_i(x) = \Phi_{m+i}(h_i(x)) \quad i = 1, \ldots, p\]

and obtain an equivalent problem

\[
\begin{align*}
\min & \quad f_0(x) \\
\text{s.t.} & \quad \tilde{f}_i(x) \leq 0 \quad i = 1, \ldots, m \\
& \quad \tilde{h}_i(x) = 0 \quad i = 1, \ldots, p
\end{align*}
\]

Example: \[\min ||Ax - b||_2^2 \quad (= f_0(x))\]

is equivalent to

\[\min (Ax - b)^T(Ax - b) \quad (= \tilde{f}_0(x))\]

Note that \(f_0(x)\) is not differentiable at \(Ax - b = 0\), but \(
\tilde{f}_0(x)\) is differentiable everywhere.

(iii) Eliminating equality constraints:

A convex problem in standard form
\[
\begin{align*}
\text{min } & f(x) \\
\text{s.t. } & f_i(x) \leq 0, \quad i = 1, \ldots, m \\
& Ax = b, \quad A \in \mathbb{R}^{p \times n}
\end{align*}
\]

is equivalent to
\[
\begin{align*}
\text{min } & f_0(Fx + x_0) \\
\text{s.t. } & f_i(Fx + x_0) \leq 0, \quad i = 1, \ldots, m
\end{align*}
\]
where \( F \) is a matrix whose range is the null space of \( A \), that is
\[Ax = b \iff x = Fz + x_0 \text{ for some } z, \quad \text{and } x_0 \text{ s.t. } Ax_0 = b.
\]

Note that \( AF = 0 \).

This transformation preserves convexity since composition of a convex function with an affine function preserves convexity.

In many cases, however, it is better (more efficient) to retain the equality constraints, especially when \( x \) is of large dimension, then equality constraints can provide sparsity or some other useful structure.

(iv) Introducing equality constraints:
\[
\begin{align*}
\text{min } & f_0(A_0x + b_0) \\
\text{s.t. } & f_i(A_0x + b_0) \leq 0, \quad i = 1, \ldots, m
\end{align*}
\]
is equivalent to
\[
\begin{align*}
\text{min } & f_0(y_0) \\
\text{s.t. } & f_i(y_i) \leq 0, \quad i = 1, \ldots, m \\
& y_i = A_i x + b_i, \quad i = 0, 1, \ldots, m
\end{align*}
\]

This problem has \( k_0 + k_m \) new variables and \( k_0 + k_m \) new constraints.
(v) **Block variables:**

Observe that \( f(x) \leq 0 \) if there is an \( s_i \geq 0 \) such that

\[
   f_i(x) + s_i = 0.
\]

Using this transformation, we obtain an equivalent problem

\[
   \min f_0(x) \\
   \text{s.t.} \quad s_i \geq 0, \quad i = 1 \ldots m \\
   f_i(x) + s_i = 0, \quad i = 1 \ldots m \\
   h_i(x) = 0, \quad i = 1 \ldots p
\]

where the variables of this new problem are \( x, s_i \) \((i=1 \ldots m)\) (or \( s \in \mathbb{R}^m \)). This new problem has \( n+m \) variables and \( m+p \) equality constraints.

(vi) **Minimizing (optimizing) over some variables:**

Observe that

\[
   \inf_{x,y} f(x,y) = \inf_x \tilde{f}(x)
\]

where \( \tilde{f}(x) = \inf_y f(x,y) \)

Thus we can partition a big problem into smaller problems

\[
   \min f_0(x_1, x_2) \\
   \text{s.t.} \quad f_i(x_1) \leq 0, \quad i = 1 \ldots m_1 \\
   \tilde{f}_i(x_2) \leq 0, \quad i = 1 \ldots m_2
\]

Since the constraints are independent (one set depends on \( x_1 \) only whereas the other set depends on \( x_2 \) only) or separable, we can first minimize over one set of variables first.

Note here \( x \in \mathbb{R}^m, \ x = (x_1, x_2) \) where \( x_1 \in \mathbb{R}^{m_1}, \ x_2 \in \mathbb{R}^{m_2} \) \( m_1 + m_2 = m \).
Define $\bar{f}(\mathbf{x}) = \inf \{ f_i(\mathbf{b}_i, \mathbf{z}) \mid f_i(z) \leq 0, i = 1, \ldots, m \}$. Then the original problem is equivalent to
\[
\min f(\mathbf{x})
\quad \text{subject to } \bar{f}_i(\mathbf{x}) \leq 0, \ i = 1, \ldots, m.
\]

Example:
\[
\min \ x_1^T P_{11} x_1 + 2 x_1^T P_{12} x_2 + x_2^T P_{22} x_2
\quad \text{subject to } \bar{f}_i(\mathbf{x}) \leq 0, \ i = 1, \ldots, m.
\]
We can analytically minimize the objective over $x_2$:
\[
\inf_{x_2} \ (x_1^T P_{11} x_1 + 2 x_1^T P_{12} x_2 + x_2^T P_{22} x_2)
\]
\[
= x_1^T (P_{11} - P_{12} P_{22}^{-1} P_{12}^T) x_1
\]
(assuming $P_{22} > 0$)
Thus, the problem is equivalent to
\[
\min \ x_1^T (P_{11} - P_{12} P_{22}^{-1} P_{12}^T) x_1
\quad \text{subject to } \bar{f}_i(\mathbf{x}) \leq 0, \ i = 1, \ldots, m.
\]

(vii) **Epigraph form:** The original problem is equivalent to
\[
\min t
\quad \text{subject to } f_0(\mathbf{x}) - t \leq 0,
\bar{f}_i(\mathbf{x}) \leq 0, \ i = 1, \ldots, m,
\mathbf{a}_i^T \mathbf{x} = \mathbf{b}_i, \ i = 1, \ldots, p.
\]

This new problem is the epigraph form and is convex, since the objective is linear and constraints are convex.

This epigraph form can be useful for theoretical analysis since it transforms any convex problem into a convex problem with linear objective.
1) Quasiconvex optimization problems:

\[
\begin{align*}
\min & \quad f_0(x) \\
\text{s.t.} & \quad f_i(x) \leq 0, \quad i = 1, \ldots, m \\
& \quad Ax = b, \quad A \in \mathbb{R}^{p \times n}
\end{align*}
\]

where \( f_0(x) \) is quasiconvex
\( f_i(x) \) is convex \( i = 1, \ldots, m \).

Quasiconvex problems can have locally optimal points that are not globally optimal.

Quasiconvex optimization can be solved through a series of convex optimizations.

Sublevel sets are convex: We can find convex representation of sublevel sets of \( f_0(x) \).

If \( f_0(x) \) is quasiconvex \( \rightarrow \exists \) a family of convex functions indexed by \( t \), \( \phi_t: \mathbb{R}^n \rightarrow \mathbb{R} \), such that

\[
f_0(x) \leq t \iff \phi_t(x) \leq 0
\]

That is, the \( t \)-sublevel set of \( f_0(x) \) is the 0-sublevel set of \( \phi_t(x) \).

A trivial example of such a representation is the indicator function

\[
\phi_t(x) = \begin{cases} 
0 & \text{if } f_0(x) \leq t \\
\infty & \text{otherwise}
\end{cases}
\]

but we can often find "nicer" functions that are differentiable.

Example: \( f_0(x) = \frac{p(x)}{q(x)} \) where \( p(x) \) is convex, \( q(x) \) concave then \( f_0(x) \) is quasiconvex. (assuming \( q(x) > 0 \), \( p(x) > 0 \).
Then \( f(x) \leq t \iff p(x) - tq(x) \leq 0 \)
and we can take \( \Phi_t(x) = p(x) - tq(x) \) for \( t \geq 0 \).

For each \( t \), \( \Phi_t(x) \) is convex in \( x \).

Quasiconvex optimization via convex feasibility problems.
Let \( p^* \) be the optimal value of the quasiconvex problem.

Consider the feasibility problem, for a fixed \( t \),

\[
\text{find } x \\
\text{s.t. } \Phi_t(x) \leq 0 \\
f_i(x) \leq 0, \quad i = 1, \ldots, m \\
Ax = b
\]

If this problem is feasible then \( p^* \leq t \), otherwise \( p^* \geq t \).

Thus we can use a bisection method for quasiconvex optimization: First obtain a lower bound \( l \) and an upper bound \( u \) on the objective function \( f(x) \).

given \( l \leq p^* \), \( u \geq p^* \) and tolerance \( \epsilon > 0 \)

repeat
1. \( t = \frac{l + u}{2} \)
2. Solve the (convex) feasibility problem
3. If problem is feasible then \( u := t \)
   else \( l := t \)

until \( u - l \leq \epsilon \).

This algorithm requires exactly \( \lceil \log_2 \frac{u-l}{\epsilon} \rceil \) iterations.

Example: \( \min (a^Tx + b)/(c^Tx + d) \) is quasiconvex optimization.

\[
\text{s.t. }Ax = b, \\
Fx \leq q, \\
c^Tx + d > 0
\]
3. Linear program (LP)

\[ \text{min } c^T x + d \]
\[ \text{st. } Gx \leq h \]
\[ Ax = b \]

All objective & constraints are affine \( \rightarrow \) linear program (LP)

The feasible set of an LP is a polyhedron \( P \).

Level sets \( \{ x \mid c^T x = \alpha \} \) are hyperplanes.

The optimal point \( x^* \) is as far as possible in \( P \) along the \( -c \) direction.

1) Standard form LP: An LP can be converted into a standard form as

\[ \text{min } c^T x \]
\[ \text{st. } Ax = b \]
\[ x \geq 0 \]

(only equality constraints and \( x \geq 0 \)).

Another form of LP is an inequality form LP:

\[ \text{min } c^T x \]
\[ \text{st. } Ax \leq b \]

Note that any general LP can be transformed into the standard form above by introducing slack variables for the inequality constraints and writing \( x = x^+ - x^- \) where \( x^+, x^- \geq 0 \).

Then the problem becomes the standard form in variables \( \{ x^+, x^-, s \} \):

\[ \text{min } c^T x^+ - c^T x^- \]
\[ \text{st. } Gx^+ - Gx^- + s = h \]
\[ Ax^+ - Ax^- = b \]
\[ x^+ \geq 0, x^- \geq 0, s \geq 0. \]
Example: (i) Chebyshev center of a polyhedron

Find the largest Euclidean ball that lies inside a polyhedron
polyhedron $P = \{ x \in \mathbb{R}^n \mid a_i^T x \leq b_i, \ i = 1, \ldots, m \}$

ball $B = \{ x + u \mid \| u \|_2 \leq r \}$

we want to find $x_c$ and the largest $r$ such that $B \subseteq P$.

To formulate the problem: For a point in $B$ to satisfy a constraint of $P$, then

$s.t. \ a_i^T (x_c + u) \leq b_i$

Thus we must have

$\max r$

$s.t. \ a_i^T x_c + r\|a_i\|_2 \leq b_i$

Thus we can formulate an LP as

Note the optimization variables are $x_c$ and $r$.

(ii) Piece-wise linear minimization:

$\min \ \max_i a_i^T x + b_i$

is equivalent to an LP

$\min t$

$s.t. \ a_i^T x + b_i \leq t, \ i = 1, \ldots, m$. The variable of this LP is $(t, x)$.

4) Linear fractional program:

This is a quasiconvex problem that can be converted into an LP.
Linear fractional prog.

\[
\begin{align*}
\text{min} & \quad \frac{c^T x + d}{e^T x + f} \\
\text{s.t.} & \quad Gx \leq h \\
& \quad Ax = b
\end{align*}
\]

\[
\begin{align*}
\min & \quad cy + dz \\
\text{s.t.} & \quad Gy \leq hz \\
& \quad Ay = bz \\
& \quad e^T y + f z = 1 \\
& \quad z \geq 0
\end{align*}
\]

\( \iff \) (variable \( x \)) (variable \( y, z \))

These two problems are equivalent since we can map

\[
y = \frac{x}{e^T x + f}, \quad z = \frac{1}{e^T x + f}
\]

4) Generalized linear fractional programming:

\[
f_0(x) = \max_{i=1,\ldots,n} \frac{c_i^T x + d_i}{e_i^T x + f_i}
\]

Example: Von Neumann model for a growing economy:

\( x \in \mathbb{R}^n \): current activity level in \( n \) sectors.

\( x^+ \in \mathbb{R}^n \): next period activity level.

\( Bx \in \mathbb{R}^m \): consumption of activity level \( x \).

\( Ax \in \mathbb{R}^m \): produce of activity level \( x \).

Then we want to maximize the growth rate \( x_i^+ / x_i \) subject to the constraint that good consumed in the next period cannot exceed good produced in the current period,

\[
\max \quad \min_{i=1,\ldots,n} \frac{x_i^+}{x_i} \quad \text{subject to} \quad Bx^+ \leq Ax
\]

\( x^+ \geq 0 \) (generalized)

4. Quadratic optimization problem (QP)

\[
\begin{align*}
\min & \quad \frac{1}{2} x^T P x + q^T x + r \\
\text{s.t.} & \quad Gx \leq h \\
& \quad Ax = b
\end{align*}
\]

\( P \in \mathbb{S}_+^n \) (PSD)

\( G \in \mathbb{R}^{m \times n}, h \in \mathbb{R}^m \)

\( A \in \mathbb{R}^{p \times n}, b \in \mathbb{R}^p \).
In QP, we minimize a convex quadratic function over a polyhedron.

**Example:** Least-square or regression

\[
\begin{align*}
\min & \quad \|Ax - b\|^2 \\
\text{s.t.} & \quad l_i \leq x_i \leq u_i, \quad i = 1 \ldots n
\end{align*}
\]

Without constraints, least-square problems have closed form solutions as \( x = A^+ b \) where \( A^+ \) is the pseudoinverse.

With even linear constraints as above, however, no closed form analytical solutions exist, but we can build efficient algorithms to solve the problem.

+) **Quadratically constrained quadratic program (QCQP):**

This is a special case of QP.

\[
\begin{align*}
\min & \quad \frac{1}{2} x^T P x + q^T x + r \\
\text{s.t.} & \quad \frac{1}{2} x^T P_i x + q^T_i x + r_i \leq 0, \quad i = 1 \ldots m \\
Ax & = b
\end{align*}
\]

The quadratic constraints define ellipsoids feasible sets if \( P_i \in S_{++}^n \) (positive definite).

+) **Second-order cone programing (SOCP)**

\[
\begin{align*}
\min & \quad f^T x \\
\text{s.t.} & \quad \|A_i x + b_i\|_2 \leq c_i x^T d_i, \quad i = 1 \ldots m \\
Fx & = g
\end{align*}
\]

where \( A_i \in \mathbb{R}^{n_i \times m} \), \( F \in \mathbb{R}^{p \times n} \)

Here the inequality constraints are cone constraints.
If $c_i = 0$ → SOCP becomes a QCQP
$A_i = 0$ → SOCP becomes an LP.

Example: Robust linear programming is an SOCP.

$$\min c^T x$$
$$\text{s.t. } a_i^T x \leq b_i, \quad i = 1, \ldots, m$$

where there can be uncertainty in $a_i$, $b_i$, and $c$.

- Deterministic approach:
  Assume $a_i \in E_i = \{ \bar{a}_i + Pu \mid \|u\|_2 \leq 1 \}$
  that is, $a_i$ lies in an "uncertainty" ellipsoid.
Then $a_i^T x \leq b_i$, $\forall a_i \in E_i$ can be expressed as

$$b_i \geq \sup \{ a_i^T x \mid a_i \in E_i \}$$

$$= \bar{a}_i^T x + \sup \{ u^T \Pi_i x \mid \|u\|_2 \leq 1 \}$$

$$= \bar{a}_i^T x + \| \Pi_i x \|_2.$$ 

Thus the robust LP problem can be expressed as

$$\min c^T x$$
$$\text{s.t. } a_i^T x + \| \Pi_i x \|_2 \leq b_i$$

- Stochastic approach:
  In a statistical framework, suppose that $a_i$ are independent
  random vectors that are Gaussian with mean $\bar{a}_i$ and covariance $\Sigma_i$.
  Formulate a robust optimization problem as

$$\min c^T x$$
$$\text{s.t. } \text{prob}(a_i^T x \leq b_i) \geq \gamma, \quad i = 1, \ldots, m$$

That is, we require that each constraint holds with
a probability exceeding $\gamma$ ($\gamma > 0.5$).

For Gaussian random $a_i$, this probability can be computed as follows.
Let $\mathbf{u}_i = \mathbf{a}_i^T \mathbf{x}$ then $\mathbf{u}_i$ is a Gaussian random variable with mean $\bar{\mathbf{u}}_i = \mathbf{a}_i^T \mathbf{x}$ and variance $\sigma_i^2 = \mathbf{x}^T \Sigma \mathbf{x}$. Now
\[
Pr(u_i \leq b_i) = Pr \left( \frac{u_i - \bar{u}_i}{\sigma_i} < \frac{b_i - \bar{u}_i}{\sigma_i} \right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{b_i - \bar{u}_i}{\sigma_i}} e^{-t^2/2} dt = \Phi \left( \frac{b_i - \bar{u}_i}{\sigma_i} \right)
\]
(\(\Phi\) is a standard Gaussian \(N(0,1)\), \(\Phi(\cdot)\) is its known CDF, thus the constraints are equivalent to
\[
\Phi \left( \frac{b_i - \bar{u}_i}{\sigma_i} \right) \geq \chi_i \quad i = 1, \ldots, m
\]
\[
\iff \frac{b_i - \bar{u}_i}{\sigma_i} \geq \phi_i \iff \bar{u}_i + \phi_i \sigma_i \leq b_i \quad i = 1, \ldots, m
\]
\[
\iff \mathbf{a}_i^T \mathbf{x} + \phi_i \sqrt{\mathbf{x}^T \mathbf{\Sigma} \mathbf{x}} \leq b_i
\]

Then the robust LP problem becomes
\[
\min c^T \mathbf{x}
\]
\[s.t. \quad \mathbf{a}_i^T \mathbf{x} + \phi_i \sqrt{\mathbf{x}^T \mathbf{\Sigma} \mathbf{x}} \leq b_i
\]
which is a SOCP problem.

Lecture 11:

5. Geometric programming:
This type of problems is not convex in their natural form but can be transformed into convex problems.
1) Monomial and polynomials:
Monomial $f: \mathbb{R}^n \to \mathbb{R}$, dom $f = \mathbb{R}_{++}^n$:
\[f(\mathbf{x}) = c_1 x_1^{a_1} \cdot \ldots \cdot x_n^{a_n} \quad c > 0\]
$a_i \in \mathbb{R}$ can be any number (positive or negative), but $c > 0$.
Note that $f(\mathbf{x}) > 0$ since dom $f = \mathbb{R}_{++}^n$. 

Let \( u_i = a_i^T x \) then \( u \) is a Gaussian random variable with mean \( \mu_i = a_i^T x \) and variance \( \sigma_i^2 = x^T \Sigma_i x \).

Now \( \Pr(u_i \leq b_i) = \Pr \left( \frac{u_i - \mu_i}{\sigma_i} < \frac{b_i - \mu_i}{\sigma_i} \right) \)

\[ = \frac{1}{2\pi} \int_{-\infty}^{b_i / \sigma_i} e^{-t^2 / 2} dt = \Phi \left( \frac{b_i - \mu_i}{\sigma_i} \right), \]

where \( \Phi(z) \) is the standard Gaussian CDF. The constraints are equivalent to

\[ \Phi \left( \frac{b_i - \mu_i}{\sigma_i} \right) \geq \eta, \quad i = 1, \ldots, m \]

\[ \iff \frac{b_i - \mu_i}{\sigma_i} \geq \Phi^{-1}(\eta) \iff \mu_i + \Phi^{-1}(\eta) \sigma_i \leq b_i, \quad i = 1, \ldots, m \]

\[ \iff a_i^T x + \Phi^{-1}(\eta) \cdot (x^T \Sigma_i x)^{1/2} \leq b_i \]

Then the robust LP problem becomes

\[ \begin{aligned}
\min & \quad c^T x \\
\text{s.t.} & \quad a_i^T x + \Phi^{-1}(\eta) \cdot \| \Sigma_i^{1/2} x \|_2 \leq b_i, \quad i = 1, \ldots, m
\end{aligned} \]

which is a SOCP problem.

---

**Lecture II:**

5. Geometric programming:

This type of problems is not convex in their natural form but can be transformed into convex problems.

- Monomial and polynomials:

  Monomial \( f : \mathbb{R}^n \rightarrow \mathbb{R} \), \( \text{dom } f = \mathbb{R}_{++}^n \):

  \[ f(x) = c x_1 + c_2 x_2 + \ldots + c_n x_n \quad \text{for } c > 0. \]

  \( a_i \in \mathbb{R} \) can be any number (\(+\) or \(-\)), but \( c > 0 \).

  Note that \( f(x) > 0 \) since \( \text{dom } f = \mathbb{R}_{++}^n \).
A polynomial is 
\[ f(\mathbf{x}) = \sum_{k=1}^{K} c_k x_1^{a_{1k}} x_2^{a_{2k}} \cdots x_n^{a_{nk}}, \quad c_k > 0 \]

- sum of monomials.

Monomials are closed under multiplication & division.
Polynomials are closed under addition, multiplication, and
non-negative scaling.

+ Geometric programming (GP).

\[ \begin{align*}
& \quad \text{min } f(\mathbf{x}) \\
& \text{s.t. } f_i(\mathbf{x}) \leq 1, \quad i = 1 \cdots m \\
& \quad h_i(\mathbf{x}) = 1, \quad i = 1 \cdots p
\end{align*} \]

where \( f_i(\mathbf{x}) \) are polynomials.
\( h_i(\mathbf{x}) \) are monomials.

The domain of the problem is \( D = \mathbb{R}_+^n \) so the constraint \( \mathbf{x} \geq 0 \) is implicit.

Example:
\[ \begin{align*}
& \quad \text{max } x/y \\
& \text{s.t. } 2 \leq x \leq 3 \\
& \quad x^2 + 3y/2 \leq \sqrt{y} \\
& \quad x/y = z^2 \\
\end{align*} \]

\[ \begin{align*}
& \quad \text{min } x^{-1} y \\
& \text{s.t. } 2x^{-1} < 1, \quad \frac{1}{3}x \leq 1 \\
& \quad x^{2y^{-1}} + 3y^{2z^{-1}} \leq 1 \\
& \quad x^{-1} y^{-1/2} z^{-2} = 1
\end{align*} \]

+ GP in convex form.

Define \( y_i = \log x_i \), \( \iff x_i = e^{y_i} \).

Then
\[ f(\mathbf{x}) = c x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n} \]
\[ = c (e^{y_1})^{a_1} (e^{y_2})^{a_2} \cdots (e^{y_n})^{a_n} \]
\[ = e^{a^T y + b} \]
\[ \implies b = \log c. \]
\[ \log f(x) = a^T y + b \quad \Rightarrow \quad y = \log x, \quad b = \log c. \]

Similarly for polynomials:

\[ f(x) = \sum_{k=1}^{K} c_k x_1 x_2 \cdots x_n^{a_{ik}} \]

\[ \Rightarrow \log f(x) = \log \left( \sum_{k=1}^{K} e^{a_{ik}x + b_k} \right) \quad \Rightarrow \quad b_k = \log c_k \]

Thus a GP can be transformed into a convex problem as:

\[ \min \log \left( \sum_{k=1}^{K} \exp \left( a_{ok}^T x + b_{ok} \right) \right) \]

s.t. \[ \log \left( \sum \exp(a_{ik}^T x + b_{ik}) \right) \leq 0, \quad i = 1, \ldots, m \]

\[ Gy + d = 0 \]

This form is also referred to as GP in convex form (versus GP in polynomial form).

If all polynomials in a GP have only one term, then the GP can be transformed into an LP.

Example: (i) Perron-Frobenius eigenvalue \( \lambda_p(A) \)

\[ A \in \mathbb{R}^{n \times n} \text{ is element-wise positive (that is } A_{ij} > 0 \text{ by} \]

Then \( \lambda_p(A) = \max_i |i(A)| \) exists.

This PF eigenvalue determines the rate of growth or decay of \( A^k \) as \( k \to \infty \), often used in control.

Given that the PF eigenvalue can be characterized as

\[ \lambda_p = \inf \{ \lambda | \lambda A_{ij} > 0 \text{ for some } i \geq 0 \} \]

Then a problem of maximizing \( \lambda_p \) of \( A(x) \) where the entries of \( A(x) \) are polynomials is a GP:

\[ \min \lambda \]

s.t. \[ \sum_{j=1}^{n} A_{ij} v_j \leq 1, \quad i = 1, \ldots, n. \]
Semi-definite programming (SDP)

\[ \begin{align*} 
\text{min} & \quad c^T x \\
\text{s.t.} & \quad \lambda_1 F_1 + \lambda_2 F_2 + \cdots + \lambda_n F_n + f_0 \leq 0 \\
& \quad A x = b \\
\end{align*} \]

where \( F_i \in S^k \) and \( A \in \mathbb{R}^{p \times n} \).

In a SDP, the inequality constraint is a linear matrix inequality (LMI).

If the matrices \( F_i \) are diagonal, then the SDP reduces to an LP.

Example: (i) LP as SDP:

\[ \begin{align*} 
\text{min} & \quad c^T x \\
\text{s.t.} & \quad A x \leq b \\
\end{align*} \]  
\((\text{LP})\)  
\[ \begin{align*} 
\text{min} & \quad c^T x \\
\text{s.t.} & \quad \text{diag}(A x - b) \leq 0 \\
\end{align*} \]  
\((\text{SDP})\)

(ii) Matrix norm minimization:

Let \( A(x) = A_0 + x_1 A_1 + \cdots + x_n A_n \), \( A_i \in \mathbb{R}^{p \times q} \).

Consider unconstrained problem:

\[ \min \quad \| A(x) \|_2 \]

Here \( \| A(x) \|_2 = \max_i \sigma_i(A) \) where \( \sigma_i \) are singular values of \( A \).

We use the fact

\[ \| A \|_2 \leq \delta \iff A^T A \leq \delta^2 I \quad (\delta > 0) \]

Then the problem becomes

\[ \min \quad \delta \]

s.t. \( A(x)^T A(x) \leq \delta^2 I \) \quad (variables \( x, \delta \))

Now using

\[ A^T A \leq \delta^2 I \iff \begin{bmatrix} tI & A \\ A^T & tI \end{bmatrix} \succeq 0 \]

we obtain the SDP

\[ \begin{align*} 
\text{min} & \quad t \\
\text{s.t.} & \quad \begin{bmatrix} tI & A \delta y \\ A \delta y^T & tI \end{bmatrix} \succeq 0 \\
\end{align*} \]  
\((\text{variables} \ x, \ t, \ \delta)\)
7) Some other examples of convex optimization problems:

- **Statistical estimation**: Covariance estimation for Gaussian data.

Suppose \( y \in \mathbb{R}^n \) is a random vector, \( y \sim \mathcal{N}(0, R) \)

\[
R = E yy^T, \quad R \in S_+^n \quad \text{is the covariance.}
\]

We want to estimate this covariance from \( N \) independent samples \( y_1, \ldots, y_N \).

\[
f_R(y) = \frac{1}{(2\pi)^{n/2} \det(R)^{1/2}} e^{-\frac{1}{2} y^T R^{-1} y}.
\]

The objective is to maximize the log-likelihood function, under some prior knowledge about \( R \).

\[
l(R) = \log f_R(y_1, y_2, \ldots, y_N)
\]

\[
= -N \frac{n}{2} - \frac{N}{2} \log \det(R) + \left( -\frac{1}{2} \sum_{k=1}^{N} y_k^T R^{-1} y_k \right)
\]

Let \( Y = \frac{1}{N} \sum_{k=1}^{N} y_k y_k^T \)

\[
\Rightarrow l(R) = -\frac{N}{2} n - \frac{N}{2} \log \det R - \frac{N}{2} \text{tr} (R^{-1} Y)
\]

Denote \( S = R^{-1} \) then \( l(R) \) is concave in \( S \).

Suppose that the prior knowledge is that there is a condition number constraint:

\[
\frac{\lambda_{\max}(R)}{\lambda_{\min}(R)} \leq K
\]

Then we can formulate a convex problem:

\[
\max \log \det S - \text{tr} (SY)
\]

s.t. \( uI \preceq S \preceq KuI \)

where the variables are \( S \in S_+^n \) and \( u > 0, u \in \mathbb{R} \).
Geometric problems: Extremal volume ellipsoids

(i) Find the minimum volume ellipsoid that contains a set \( C \):

\[
E = \{ x \mid \| A x + b \|_2 \leq 1 \ \forall \ x \in C \}.
\]

Assuming that \( A \in S_+^n \) which is invertible, then the volume of the ellipsoid is proportional to \( \det A^{-1} \).

Then the minimum volume can be found through an opt prob.

\[
\min \ \log \det A^{-1} \\
\text{s.t.} \ \sup_{u \in C} \| A u + b \|_2 \leq 1
\]

This problem is convex, but evaluating the constraint can be hard (for general \( C \)).

For a finite set \( C = \{ x_1, \ldots, x_m \} \):

\[
\min \ \log \det A^{-1} \\
\text{s.t.} \ \| A x_i + b \|_2 \leq 1, \ i = 1, \ldots, m
\]

→ this problem solves for the Löwner-John ellipsoid for the polyhedron conv \( \{ x_1, \ldots, x_m \} \). (Consider this polyhedron)

(ii) Find the maximum volume ellipsoid contained in a set:

For polyhedron

\[
C = \{ x \mid a_i^T x \leq b_i, \ i = 1, \ldots, m \}
\]

want to find the max ellipsoid

\[
E = \{ Bu + d \mid \| u \|_2 \leq 1 \}
\]

that is inside set \( C \).

Constraint:

\[
\sup_{\| u \|_2 \leq 1} a_i^T (Bu + d) \leq b_i \iff \| Ba_i \|_2 + a_i^T d \leq b_i
\]

Thus we can formulate

\[
\min \ \log \det B^{-1} \\
\text{s.t.} \ \| Ba_i \|_2 + a_i^T d \leq b_i
\]
Generalized polynomials and GP:

\[ f_i : \mathbb{R}^n \rightarrow \mathbb{R} \quad \text{polynomials}, \quad i = 1 \ldots k \]

\[ \Phi : \mathbb{R}^k \rightarrow \mathbb{R} \quad \text{polynomial with non-negative coefficients, but can have fractional exponents} \]

Then \[ h(x) = \Phi(f_1(x), \ldots, f_k(x)) \]

is a generalized polynomial.

Example:

\[ \Phi_1(z_1, z_2) = 3z_1^2z_2 + 2z_1 + 3z_2^3 \]

\[ \rightarrow h = 3f_1^2f_2 + 2f_1 + f_2^3 \quad \text{is a polynomial} \]

\[ \Phi_2(z_1, z_2) = 2z_1^{0.3}z_2^{1.2} + z_1z_2^{0.5} + 2 \]

\[ \rightarrow h = 2f_1^{0.3}f_2^{1.2} + f_1f_2^{0.5} + 2 \quad \text{is a generalized polynomial} \]

Generalized GP:

\[ \min h_0(x) \]

s.t. \[ h_i(x) \leq 1 \quad i = 1 \ldots m \]

\[ g_i(x) = 1 \quad i = 1 \ldots p \]

\[ g_i(x) : \text{monomials}, \quad h_i(x) : \text{generalized polynomial} \]

This problem can be transformed into a GP by introducing new variables.