3341.454: Mathematical and Numerical Optimization **SNU, Spring 2026**
Lecture #3: Linear Programming March 11, 2026
Lecturer: Dabeen Lee

In this lecture, we consider

- Convex programming hierarchy,
- Introduction to linear programming,
- Optimization problems represented as linear programs.

1 Convex Programming Hierarchy

On top of the applications we studied, there are many interesting "classes" of convex optimization problems. In this lecture, we consider them in this section, and specifically, we discuss

- Linear programming (LP),
- Quadratic programming (QP),
- Semidefinite programming (SDP),
- Conic programming,
- Second-order cone programming (SOCP).

In fact, these problems are closely related, and in fact, the problem classes form a hierarchy described in Figure 3.1

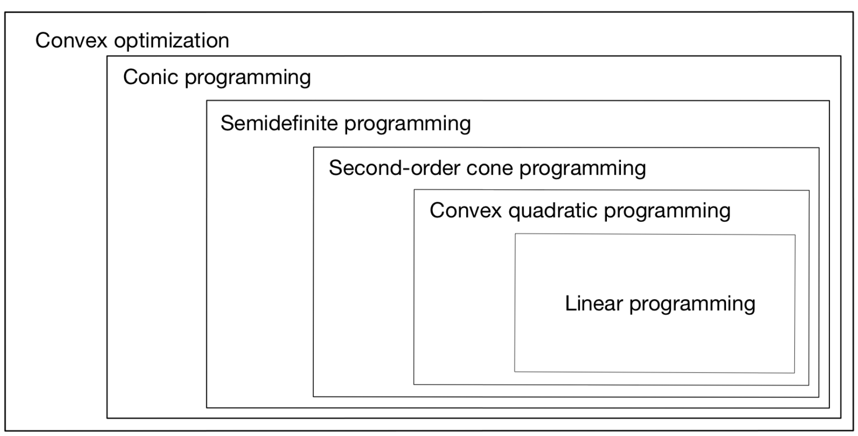

Figure 3.1: Hierarchy of classes of convex optimization problems

2 Linear Programming

A linear program (LP) is an optimization problem of the following form.

$$\begin{aligned} \text{minimize} \quad & c^\top x \\ \text{subject to} \quad & Ax \geq b \end{aligned} \tag{P}$$

where $c^\top x$ is the linear objective function and $Ax \geq b$ is the system of linear constraints $a_1^\top x \geq b_1, \ldots, a_n^\top x \geq b_n$, i.e., $a_1^\top, \ldots, a_n^\top$ are the rows of A and $b = (b_1, \ldots, b_n)^\top$. Note that the feasible region $P = \{x \in \mathbb{R}^d : Ax \geq b\}$ is a polyhedron, the intersection of half-spaces $\{x \in \mathbb{R}^d : a_i^\top x \geq b_i\}$ for $i = 1, \ldots, n$. Hence, the linear program is the problem of finding a point in a polyhedron that minimizes a linear function. Figure 3.2 describes a linear program with 2 variables.

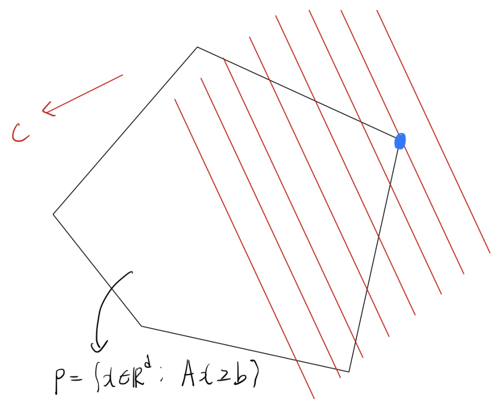

Figure 3.2: Geometric picture of a linear program

Example 3.1 (Production planning)**.** We consider the following production planning problem. A company produces d products. Each product $j \in [d]$ requires a_{ij} units of resource $i \in [m]$ to produce one unit of product j. The company has b_i units of resource $i \in [m]$ available. The profit for producing one unit of product j is p_j. The company wants to maximize the total profit. This problem can be formulated as the following linear program.

$$\begin{aligned} \text{maximize} \quad & p^\top x \\ \text{subject to} \quad & Ax \leq b \\ & x \geq 0 \end{aligned}$$

where x_j is the number of units of product j to produce, $A = (a_{ij})_{i \in [m], j \in [d]}$ is the resource consumption matrix, $b = (b_1, \ldots, b_m)^\top$ is the resource availability vector, and $p = (p_1, \ldots, p_d)^\top$ is the profit vector.

Example 3.2 (Transportation problem)**.** We consider the following transportation problem. There are m factories and d warehouses. Each factory $i \in [m]$ has a supply of s_i units of goods, and each warehouse $j \in [d]$ has a demand of d_j units of goods. The cost of transporting one unit of goods from factory i to warehouse j is c_{ij}. The goal is to find a transportation plan that minimizes the total transportation cost while satisfying the supply and demand constraints. This problem can be formulated as the following linear program.

$$\begin{aligned}
\text{minimize} \quad & \sum_{i=1}^{m}\sum_{j=1}^{d} c_{ij}x_{ij} \\
\text{subject to} \quad & \sum_{j=1}^{d} x_{ij} \le s_i, \quad i = 1,\ldots,m \\
& \sum_{i=1}^{m} x_{ij} \ge d_j, \quad j = 1,\ldots,d \\
& x_{ij} \ge 0, \quad i = 1,\ldots,m, j = 1,\ldots,d
\end{aligned}$$

where x_{ij} is the amount of goods transported from factory i to warehouse j.

3 Optimization Problems Represented as Linear Programs

In general, what type of decision-making problems can be modeled as linear programs? Recall that an optimization problem is of the form

$$\begin{aligned}
\min \quad & f(x) \\
\text{s.t.} \quad & g_i(x) \le 0, \quad i \in [m], \\
& x \in \mathbb{R}^d.
\end{aligned}$$

Here, if the objective function f and the constraint functions $g_1, \ldots, g_m$ are linear, then the optimization problem is a linear program. In fact, **even when the functions are non-linear, we may still be able to represent the problem as a linear program**.

Consider the following problem. Let $V = \left\{v^1, \ldots, v^n\right\} \subseteq \mathbb{R}^d$ be a set of vectors in a cluster. The problem is to determine the center of the cluster. Here, the center is a point, from which the distance to the set of n vectors is minimized. Depending on which distance metric is used, we may obtain different centers.

Based on the vector norms, we will define the distance between a point x and the set V, denoted as $d(x)$, and we set the center of the cluster by solving

$$\min_x \quad d(x).$$

Now we consider the following four distance functions.

1. The sum of the ℓ_1-distance between x and individual vector v^i for $i \in [n]$.

$$d(x) = \sum_{i \in [n]} \left\| x - v^i \right\|_1.$$

2. The sum of the ℓ_∞-distance between x and individual vector v^i for $i \in [n]$.

$$d(x) = \sum_{i \in [n]} \left\| x - v^i \right\|_\infty.$$

3. The maximum ℓ_1-distance between x and individual vector v^i for $i \in [n]$.

$$d(x) = \max_{i \in [n]} \left\| x - v^i \right\|_1.$$

4. The maximum ℓ_∞-distance between x and individual vector v^i for $i \in [n]$.

$$d(x) = \max_{i\in[n]} \left\|x - v^i\right\|_\infty .$$

Here, as the vector norms are not linear, the distance functions are not linear either. Nevertheless, we will see that the problem for each of the four distance functions can be rewritten as a linear program.

3.1 Linearly Representable Functions

Recall that the **epigraph** of a function $f : \mathbb{R}^d \to \mathbb{R}$ is defined as

$$\mathrm{epi}(f) = \{(x,t) \in \mathbb{R}^d \times \mathbb{R} : \ f(x) \le t\} \subseteq \mathbb{R}^{d+1}.$$

A function $f : \mathbb{R}^d \to \mathbb{R}$ is **linearly representable** if its epigraph $\mathrm{epi}(f)$ is represented with a **finite** system of linear inequalities, i.e.,

$$\mathrm{epi}(f) = \left\{(x,t) \in \mathbb{R}^d \times \mathbb{R} : \ \exists y \in \mathbb{R}^p \text{ s.t. } Ax + Dy + ht \le r\right\}$$

for some $A \in \mathbb{R}^{\ell\times d}$, $D \in \mathbb{R}^{\ell\times p}$, and $h, r \in \mathbb{R}^\ell$. Here, the variables y in the linear representation are called **auxiliary variables**. We want variables (x, y, t) to satisfy the linear inequalities

$$Ax + Dy + ht \le r,$$

but $\mathrm{epi}(f)$ collects the points that has the (x,t) part.

Theorem 3.3. *The optimization problem*

$$\min\left\{f(x) : \ g_i(x) \le 0, \ \forall i \in [m]\right\}$$

can be reformulated as a linear program if the objective function and the constraint functions are linearly representable.

Proof. The objective is to minimize $f(x)$. We may reformulate this by introducing an auxiliary variable $t \in \mathbb{R}$. To be specific, minimizing $f(x)$ is equivalent to minimizing t subject to an additional constraint $f(x) \le t$. Therefore, the optimization is equivalent to

$$\min\left\{t : \ f(x) \le t, \ g_i(x) \le 0, \ \forall i \in [m]\right\}.$$

Using the notion of epigraphs, we can equivalently write this as

$$\min\left\{t : \ (x,t) \in \mathrm{epi}(f), \ (x,0) \in \mathrm{epi}(g_i), \ \forall i \in [m]\right\}.$$

Since f is linearly representable,

$$\mathrm{epi}(f) = \left\{(x,t) \in \mathbb{R}^d \times \mathbb{R} : \ \exists y \in \mathbb{R}^p \text{ s.t. } Ax + Dy + ht \le r\right\}$$

for some $A \in \mathbb{R}^{\ell\times d}$, $D \in \mathbb{R}^{\ell\times p}$, and $h, r \in \mathbb{R}^\ell$. Moreover, since $g_1, \ldots, g_m$ are linearly representable,

$$\mathrm{epi}(g_i) = \left\{(x,t) \in \mathbb{R}^d \times \mathbb{R} : \ \exists z^i \in \mathbb{R}^{q_i} \text{ s.t. } A^i x + D^i z^i + h^i t \le r^i\right\}$$

for some $A^i \in \mathbb{R}^{\ell_i \times d}$, $D^i \in \mathbb{R}^{\ell_i \times p}$, and $h^i, r^i \in \mathbb{R}^{\ell_i}$. Therefore, the optimization problem is equivalent to

$$\begin{aligned}
\min \quad & t \\
\text{s.t.} \quad & Ax + Dy + ht \leq r, \\
& A^i x + D^i z^i \leq r^i, \quad i \in [m], \\
& x \in \mathbb{R}^d,\ t \in \mathbb{R},\ y \in \mathbb{R}^p,\ z_i \in \mathbb{R}^{q_i},\ i \in [m],
\end{aligned}$$

which is a linear program. □

By definition, any linear function is linearly representable. In fact, there exist functions that are non-linear but linearly representable.

Example 3.4. $f(x_1, x_2, x_3) = \max\{x_1,\ 2x_2 + x_3,\ 2x_3 - x_1\}$. Its epigraph is given by

$$\{(x_1, x_2, x_3, t) :\ \max\{x_1,\ 2x_2 + x_3,\ 2x_3 - x_1\} \leq t\}$$

Note that $\max\{x_1,\ 2x_2 + x_3,\ 2x_3 - x_1\} \leq t$ is equivalent to

$$x_1, \leq t,\ 2x_2 + x_3 \leq t,\ 2x_3 - x_1 \leq t.$$

Therefore, it follows that

$$\text{epi}(f) = \{(x_1, x_2, x_3, t) :\ x_1 \leq t,\ 2x_2 + x_3 \leq t,\ 2x_3 - x_1 \leq t\},$$

so f is linearly representable.

3.2 Projection

For the linear representation of epi(f), we have auxiliary variables y. Here, what does it mean by having auxiliary variables y? To answer this, we discuss the concept of **projection**. Let $C \subseteq \mathbb{R}^d \times \mathbb{R}^p$ be a set in the space $\mathbb{R}^d \times \mathbb{R}^p$. For convenience, we represent each point in C as (x, y) where $x \in \mathbb{R}^d$ and $y \in \mathbb{R}^p$. Then the projection of C onto the space of x part[1], which is $\mathbb{R}^d$, is given by

$$\text{proj}_x(C) = \left\{x \in \mathbb{R}^d :\ \exists y \in \mathbb{R}^p \text{ s.t. } (x, y) \in C\right\}.$$

We also refer to this operation as **projecting out the y variables**. Figure 3.3 illustrates the projection of a set in $\mathbb{R} \times \mathbb{R}$ to $\mathbb{R}$.

Getting back to epi(f), we take the set of vectors that have the (x, y, t) part and obtain

$$P = \left\{(x, y, t) \in \mathbb{R}^d \times \mathbb{R}^p \times \mathbb{R} :\ Ax + Dy + ht \leq r\right\}.$$

Then epi(f) is the projection of P onto the (x, t)-space. Here, P is a polyhedron, and epi(f) is the projection of a polyhedron. Therefore, we can also say that a function is linearly representable if its epigraph is the projection of a polyhedron.

In fact, the projection of a polyhedron is still a polyhedron. This is a consequence of the following result, known as **Fourier-Motzkin elimination**.

Theorem 3.5 (Fourier-Motzkin elimination)**.** *The projection of a polyhedron is still a polyhedron.*

[1] This may not be the most conventional terminology.

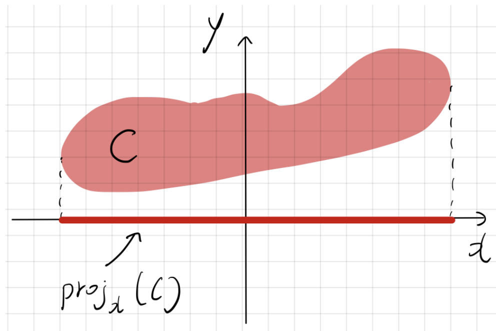

Figure 3.3: Projection of a set in $\mathbb{R}^2$ to $\mathbb{R}$

Proof. Let $P = \{(x, y) \in \mathbb{R}^d \times \mathbb{R}^p : \ Ax + Dy \leq r\}$ be a polyhedron. We want to show that $\text{proj}_x(P)$ is a polyhedron. Let $A = (a_{ij})_{i\in[\ell],j\in[d]}$ and $D = (d_{ij})_{i\in[\ell],j\in[p]}$. We partition the inequalities in $Ax + Dy \leq r$ into three groups:

- Group 1: $d_{i1} > 0$ for $i \in I_1$,
- Group 2: $d_{i1} < 0$ for $i \in I_2$,
- Group 3: $d_{i1} = 0$ for $i \in I_3$.

Then we can rewrite the inequalities as follows.

$$
\begin{aligned}
a_{ij}x_j + \sum_{j=2}^{d} a_{ij}x_j + d_{i1}y_1 + \sum_{j=2}^{p} d_{ij}y_j \leq r_i, \quad i \in I_1, \\
a_{ij}x_j + \sum_{j=2}^{d} a_{ij}x_j + d_{i1}y_1 + \sum_{j=2}^{p} d_{ij}y_j \leq r_i, \quad i \in I_2, \\
a_{ij}x_j + \sum_{j=2}^{d} a_{ij}x_j + \sum_{j=2}^{p} d_{ij}y_j \leq r_i, \quad i \in I_3.
\end{aligned}
$$

The projection of P onto the x-space is given by

$$
\text{proj}_x(P) = \{x : \ \exists y_1, \ldots, y_p \text{ s.t. } Ax + Dy \leq r\},
$$

where $y_1, \ldots, y_p$ are the auxiliary variables. We can eliminate y_1 by combining the inequalities in Group 1 and Group 2. Specifically, for any $i \in I_1$ and $j \in I_2$, we can multiply the inequality in Group 1 by $-d_{j1}$ and the inequality in Group 2 by d_{i1}, and then add them together to eliminate y_1. This gives us a new inequality that does not involve y_1. After eliminating y_1, we can repeat the same process to eliminate y_2, and so on, until all auxiliary variables are eliminated. The resulting set of inequalities will define $\text{proj}_x(P)$, which is a polyhedron. $\square$

3.3 Common Linearly Representable Functions

In the previous section, we saw some linearly representable functions. In this section, we conver more linearly representable functions.

Absolute value Let $f(x) = |x|$ for $x \in \mathbb{R}$. Note that $|x| \leq t$ if and only if $-t \leq x \leq t$. Therefore,

$$\mathrm{epi}(f) = \{(x,t) \in \mathbb{R} \times \mathbb{R} : \ -t \leq x \leq t\}.$$

Convex piecewise linear functions Let $f(x) = \max_{i \in [n]} \left\{c_i^\top x + d_i\right\}$. Note that $f(x) = \max_{i\in[n]} \left\{c_i^\top x + d_i\right\} \leq t$ if and only if $c_i^\top x + d_i \leq t$ for $i \in [n]$. Then

$$\mathrm{epi}(f) = \left\{(x,t) \in \mathbb{R}^d \times \mathbb{R} : \ c_i^\top x + d_i \leq t, \quad \forall i \in [n]\right\}.$$

ℓ_1-norm Let $f(x) = \|x\|_1 = \sum_{j\in[d]} |x_j|$. Note that $\|x\|_1 \leq t$ if and only if there exist $s_1, \ldots, s_d \in \mathbb{R}$ such that

$$|x_j| \leq s_j, \ \forall j \in [d], \quad \sum_{j\in[d]} s_j \leq t.$$

This is equivalent to

$$-s_j \leq x_j \leq s_j, \ \forall j \in [d], \quad \sum_{j\in[d]} s_j \leq t.$$

ℓ_∞-norm Let $f(x) = \|x\|_\infty = \max_{j\in[d]} |x_j|$. Note that $\|x\|_\infty \leq t$ if and only if

$$|x_j| \leq t \ \forall j \in [d]$$

This is equivalent to

$$-t \leq x_j \leq t, \ \forall j \in [d].$$

In fact, there are some operations that preserve linear representability, with which we can build more complex linearly representable functions.

Theorem 3.6. *Let $f_1(x), \ldots, f_n(x)$ be linearly representable functions. Then the following statements hold.*

1. *For any non-negative scalars $\alpha_1, \ldots, \alpha_n$,*
$$f(x) = \sum_{i\in[n]} \alpha_i f_i(x)$$
is linearly representable,
2. *The point-wise maximum of $f_1, \ldots, f_n$, given by*
$$f(x) = \max_{i\in[n]} f_i(x)$$
is linearly representable.

Proof. For the first part,

$$\begin{aligned}
\mathrm{epi}(f) &= \left\{(x,t) \in \mathbb{R}^d \times \mathbb{R} : \ \sum_{i\in[n]} \alpha_i f_i(x) \leq t\right\} \\
&= \left\{(x,t) \in \mathbb{R}^d \times \mathbb{R} : \ \exists s \in \mathbb{R}^n \text{ s.t. } \sum_{i\in[n]} \alpha_i s_i \leq t, \ f_i(x) \leq s_i, \ \forall i \in [n]\right\} \\
&= \left\{(x,t) \in \mathbb{R}^d \times \mathbb{R} : \ \exists s \in \mathbb{R}^n \text{ s.t. } \sum_{i\in[n]} \alpha_i s_i \leq t, \ (x, s_i) \in \mathrm{epi}(f_i), \ \forall i \in [n]\right\}.
\end{aligned}$$

For the second part,

$$
\begin{aligned}
\operatorname{epi}(f) &= \left\{(x,t) \in \mathbb{R}^d \times \mathbb{R} : \ \max_{i\in[n]} f_i(x) \le t\right\} \\
&= \left\{(x,t) \in \mathbb{R}^d \times \mathbb{R} : \ f_i(x) \le t, \ \forall i \in [n]\right\} \\
&= \left\{(x,t) \in \mathbb{R}^d \times \mathbb{R} : \ (x,t) \in \operatorname{epi}(f_i), \ \forall i \in [n]\right\}.
\end{aligned}
$$

In both cases, $\operatorname{epi}(f)$ is represented by a finite system of linear inequalities because $f_1, \ldots, f_n$ are linearly representable. $\square$