Types of LPs

CSCI 4113/6101
Instructor: Norbert Zeh
September 10, 2025

So far, we have focused on discussing how to model some optimization problems using LPs. More examples will follow in subsequent chapters—linear programming will be used as a tool to express optimization problems throughout this course. The next chapter is concerned with discussing the most widely used algorithm for solving LPs, the Simplex Algorithm. This algorithm assumes that its input is in a particular form. This topic introduces two standard representations of linear programs, the canonical form and the standard form; the latter, sometimes also called **slack form**, is the one the Simplex Algorithm works with.[1]

1 Canonical Form

An LP in **canonical form** consists of an n-element row vector

$$c = (c_1, \ldots, c_n),$$

an m-element column vector

$$b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix},$$

and an $m \times n$ matrix

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix}.$$

[1] To make things perfectly confusing, texts that refer to standard form as slack form often also refer to canonical form as standard form. I follow the terminology used by Papadimitriou and Steiglitz (1982) here because the interpretation of some variables in the standard form as "slack variables", which justifies the name "slack form", makes sense only immediately after converting an LP in canonical form into one in standard form.

A, b, and c represent the following LP:

$$\begin{aligned} \text{Maximize } & \sum_{j=1}^{n} c_j x_j \\ \text{s.t. } & \sum_{j=1}^{n} a_{ij} x_j \le b_i \quad \forall i \in [m] \\ & x_j \ge 0 \quad \forall j \in [n] \end{aligned} \tag{1}$$

The last n constraints are called **non-negativity constraints**.

In other words, all variables in an LP in canonical form are required to be non-negative, all constraints are upper bound constraints—constant upper bounds on linear combinations of the variables—and the objective is to maximize the objective function.

If we define that $x \le y$ for two m-element vectors if and only if $x_i \le y_i$ for all $1 \le i \le m$, and we write 0 to denote the 0-vector $(0, \dots, 0)^T$, then this LP can be written more compactly as[2]

$$\begin{aligned} \text{Maximize } & cx \\ \text{s.t. } & Ax \le b \\ & x \ge 0, \end{aligned}$$

where cx is the inner product of the two vectors c and x and Ax is the usual matrix-vector product. (Throughout the discussion of linear programs in this course, we treat x as a column vector.)

Throughout most of this course, we call two LPs **equivalent** if they have the same variables and the same set of feasible solutions, and their objective functions assign the same value to each feasible solution. In the following two lemmas, we use the more general notion of equivalence that considers two LPs P and P' over variables $x = (x_1, \dots, x_n)^T$ and $y = (y_1, \dots, y_r)^T$ equivalent if there exist two linear functions $\phi : \mathbb{R}^n \to \mathbb{R}^r$ and $\psi : \mathbb{R}^r \to \mathbb{R}^n$ such that

- $\hat{x}$ is a feasible solution of P if and only if $\phi(\hat{x})$ is a feasible solution of P',
- $\hat{y}$ is a feasible solution of P' if and only if $\psi(\hat{y})$ is a feasible solution of P,
- If $\hat{x}$ is an optimal solution of P, then $\phi(\hat{x})$ is an optimal solution of P', and
- If $\hat{y}$ is an optimal solution of P', then $\psi(\hat{y})$ is an optimal solution of P.

Equivalence of LPs is useful because it allows us to obtain an optimal solution of an LP by solving any of its equivalent LPs.

Proposition 1. *Every linear program can be transformed into an equivalent linear program in canonical form.*

Proof. An LP P may not be in canonical form for four reasons:

- It may be a minimization LP.
- Some of its constraints may be equality constraints.
- Some of its inequality constraints may be lower bounds, as opposed to upper bounds.

[2] Recall that A^T denotes the transpose of the matrix A. A column vector is simply an $m \times 1$ matrix. A row vector is simply a $1 \times n$ matrix. Here, transposing the row vector $(0, \dots, 0)$ produces a column vector.

- Some variables may not have non-negativity constraints.

We define a sequence of LPs $P = P_0, P_1, P_2, P_3, P_4 = P'$ such that P' is in canonical form and for all $1 \leq i \leq 4$, P_{i-1} and P_i are equivalent. Thus, P' is an LP in canonical form that is equivalent to P.

From minimization LP to maximization LP. First we construct a maximization LP P_1 that is equivalent to $P_0 = P$. If P_0 is itself a maximization LP, then $P_1 = P_0$. Otherwise, the objective of P_0 is to minimize cx. This is the same as maximizing $-cx$. Thus, P_1 has the objective function $-cx$ and asks us to maximize it. P_0 and P_1 have the same variables and the same constraints, so they have the same set of feasible solutions. As just observed, any feasible solution that minimizes cx also maximizes $-cx$. Thus, P_0 and P_1 are equivalent.

From equalities to inequalities. The next LP P_2 we construct from P_1 is a maximization LP without equality constraints. Clearly,

$$\sum_{j=1}^{n} a_{ij} x_j = b_i$$

if and only if

$$\sum_{j=1}^{n} a_{ij} x_j \geq b_i \text{ and}$$
$$\sum_{j=1}^{n} a_{ij} x_j \leq b_i.$$

Thus, we construct P_2 by replacing every equality constraint in P_1 with these two corresponding inequality constraints. P_1 and P_2 have the same variables and the same objective function, and every inequality constraint in P_1 is also a constraint in P_2. As just observed, any solution $\hat{x} = (\hat{x}_1, \ldots, \hat{x}_n)$ satisfies an equality constraint in P_1 if and only if it satisfies the two corresponding inequality constraints in P_2. Thus, P_1 and P_2 have the same set of feasible solutions and the same objective function. They are therefore equivalent.

From lower bounds to upper bounds. The next LP P_3 we construct from P_2 is a maximization LP with only upper bound constraints, apart from constraints of the form $x_j \geq 0$ or $x_j \leq 0$. This transformation is similar to the transformation from a minimization LP to a maximization LP: Every lower bound constraint

$$\sum_{j=1}^{n} a_{ij} x_j \geq b_i$$

in P_2 is satisfied if and only if the upper bound constraint

$$\sum_{j=1}^{n} -a_{ij} x_j \leq -b_i$$

is satisfied, so we replace it with this constraint in P_3. Since P_2 and P_3 have the same variables and the same objective function, and it is easy to see that $\hat{x} = (\hat{x}_1, \ldots, \hat{x}_n)$ is a feasible solution of P_2 if and only if it is a feasible solution of P_3, P_2 and P_3 are equivalent.

Introducing non-negativity constraints. P_3 is in canonical form, except that there may be variables that are constrained to be non-positive ($x_j \le 0$) and there may be variables that are unconstrained, that is, that are neither required to be non-negative nor to be non-positive. Let X_+, X_-, and $X_\pm$ be the sets of variables in P_3 that are constrained to be non-negative, constrained to be non-positive, and unconstrained, respectively. We construct a new set of variables

$$Y = \{y_j \mid x_j \in X_+ \cup X_\pm\} \cup \{z_j \mid x_j \in X_- \cup X_\pm\}.$$

P_4 is obtained from P_3 by adding the constraints $y_j \ge 0$ and $z_j \ge 0$, for all $y_j, z_j \in Y$, and performing the following substitutions:

- Replace every occurrence of every variable $x_j \in X_+$ in P_3 with y_j,
- Replace every occurrence of every variable $x_j \in X_-$ in P_3 with $-z_j$, and
- Replace every occurreence of every variable $x_j \in X_\pm$ in P_3 with $y_j - z_j$.

The linear transformations ϕ and ψ that establish the equivalence between P_3 and P_4 are defined as follows:

$$\phi(\hat{x}) = (\hat{y}, \hat{z}),$$

where $\hat{y} = (\hat{y}_j)_{x_j \in X_+ \cup X_\pm}$ and $\hat{z} = (\hat{z}_j)_{x_j \in X_- \cup X_\pm}$, such that

$$\begin{aligned} \hat{y}_j &= \max(\hat{x}_j, 0) && \forall x_j \in X_+ \cup X_\pm, \\ \hat{z}_j &= \max(-\hat{x}_j, 0) && \forall x_j \in X_- \cup X_\pm, \end{aligned}$$

and

$$\psi(\hat{y}, \hat{z}) = \hat{x},$$

where

$$\begin{aligned} \hat{x}_j &= \hat{y}_j && \forall x_j \in X_+, \\ \hat{x}_j &= -\hat{z}_j && \forall x_j \in X_-, \\ \hat{x}_j &= \hat{y}_j - \hat{z}_j && \forall x_j \in X_\pm. \end{aligned}$$

Given a feasible solution $\hat{x}$ of P_3, we have $\phi(\hat{x}) = (\hat{y}, \hat{z})$, where the definition of ϕ explicitly ensures that $\hat{y}_j \ge 0$ and $\hat{z}_j \ge 0$, for all $y_j, z_j \in Y$, that is, all variables in Y satisfy the non-negativity constraints in P_4. Given the replacement of variables performed to obtain P_4 from P_3, $(\hat{x}, \hat{y})$ satisfies all other constraints in P_4 if we can prove that

(i) $\hat{y}_j = \hat{x}_j$, for all $x_j \in X_+$,
(ii) $-\hat{z}_j = \hat{x}_j$, for all $x_j \in X_-$, and
(iii) $\hat{y}_j - \hat{z}_j = \hat{x}_j$, for all $x_j \in X_\pm$.

(i) For every $x_j \in X_+$, P_3 contains the constraint $x_j \ge 0$. Since $\hat{x}$ is a feasible solution of P_3, this implies that $\hat{x}_j \ge 0$, so $\hat{y}_j = \max(\hat{x}_j, 0) = \hat{x}_j$.

(ii) For every $x_j \in X_-$, P_3 contains the constraint $x_j \le 0$. Since $\hat{x}$ is a feasible solution of P_3, this implies that $\hat{x}_j \le 0$, so $\hat{z}_j = \max(-\hat{x}_j, 0) = -\hat{x}_j$.

(iii) For every real number x, it is easily verified that $x = \max(x, 0) - \max(-x, 0)$. Thus, for every $x_j \in X_\pm$, we have $\hat{y}_j - \hat{z}_j = \max(\hat{x}_j, 0) - \max(-\hat{x}_j, 0) = \hat{x}_j$.

Since the same replacements of variables are performed in the objective function of P_3 to obtain the objective function of P_4, these three properties also immpdiately imply that $\hat{x}$ and $(\hat{y}, \hat{z})$ have the same objective function value as solutions of P_3 and P_4, respectively.

Given a feasible solution $(\hat{y}, \hat{z})$ of P_4, let $\hat{x} = \psi(\hat{y}, \hat{z})$. Then all values in $(\hat{y}, \hat{z})$ are non-negative due to the non-negativity constraints in P_4. Thus, $\hat{x}_j = \hat{y}_j \ge 0$, for all $x_j \in X_+$, that is, all variables of P_3 that are required to be non-negative have non-negative values. Similarly, $\hat{x}_j = -\hat{z}_j \le 0$, for all $x_j \in X_-$, that is, all variables in P_3 that are required to be non-positive have non-positive values. Given that $(\hat{y}, \hat{z})$ satisifies all the constraints in P_4 and given the replacement of variables performed to obtain P_4 from P_3, the definition of $\hat{x}$ immediately implies that $\hat{x}$ satisfies all remaining constraints in P_3, and that $\hat{x}$ and $(\hat{y}, \hat{z})$ have the same objective function value as solutions of P_3 and P_4, respectively. This finishes the proof that P_3 and P_4 are equivalent.

Since we have shown that $P_0, \dots, P_4$ are all equivalent and P_4 is in canonical form, the proposition holds. □

Example 2. Here is an illustration of the transformation from Prop. 1. It uses a slightly modified version of an LP that arises in modelling minimum-cost flow problems. This LP describes a flow through a graph G whose edges are partitioned into two subsets E_b and E_u. The LP has one variable p_v per vertex $v \in V$ and one variable $s_{u,v}$ for every edge $(u, v) \in E_b$. There is no variable corresponding to any edge in E_u:

$$\begin{aligned} \text{Minimize} \sum_{(u,v)\in E_b} c_{u,v} s_{u,v} - \sum_{u\in V} b_u p_u & \\ \text{s.t. } p_u - p_v - s_{u,v} \le q_{u,v} & \quad \forall (u,v) \in E_b \\ p_v - p_u \ge q_{u,v} & \quad \forall (u,v) \in E_u \\ s_{u,v} \ge 0 & \quad \forall (u,v) \in E_b \end{aligned}$$

First we turn the LP into a maximization LP:

$$\text{Maximize} \sum_{u\in V} b_u p_u - \sum_{(u,v)\in E_b} c_{u,v} s_{u,v}$$

There are no equality constraints, so the second transformation is not needed. Next, we turn lower bounds into upper bounds:

$$\begin{aligned} p_u - p_v - s_{u,v} \le q_{u,v} & \quad \forall (u,v) \in E_b \\ p_u - p_v \le -q_{u,v} & \quad \forall (u,v) \in E_u \end{aligned}$$

Finally, observe that all variables $s_{u,v}$ for $(u, v) \in E_b$ are already constrained to be non-negative, but the variables p_u for $u \in V$ are unconstrained. Thus, we replace every variable p_u with two variables p'_u and

p''_u to obtain the final LP in canonical form:

$$\begin{aligned}
\text{Maximize } \sum_{u \in V} b_u p'_u - \sum_{u \in V} b_u p''_u - \sum_{(u,v) \in E_b} c_{u,v} s_{u,v} & \\
\text{s.t. } p'_u - p''_u - p'_v + p''_v - s_{u,v} \le q_{u,v} & \quad \forall (u,v) \in E_b \\
p'_u - p''_u - p'_v + p''_v \le -q_{u,v} & \quad \forall (u,v) \in E_u \\
p'_u \ge 0 & \quad \forall u \in V \\
p''_u \ge 0 & \quad \forall u \in V \\
s_{u,v} \ge 0 & \quad \forall (u,v) \in E_b
\end{aligned} \tag{2}$$

1.1 Standard Form

The Simplex Algorithm expects its input in **standard form**:

$$\begin{aligned}
\text{Maximize } cx + d & \\
\text{s.t. } Ax + y = b & \\
x \ge 0 & \\
y \ge 0, &
\end{aligned} \tag{3}$$

where c is an n-element row vector, d is an arbitrary real number, b is an m-element column vector, A is an $m \times n$-matrix, and y is a column vector of m variables. In other words, just as for an LP in canonical form, the objective is to maximize the objective function and all variables are constrained to be non-negative. However, whereas all constraints in an LP in canonical form are required to be upper bound constraints, the constrains in an LP in standard form must be equality constraints. In addition, there must be a one-to-one correspondence between the variables in y and the constraints, in the sense that every variable in y appears with coefficient 1 in its corresponding constraint and with coefficient 0 everywhere else, even in the objective function.

The elements of y are called **basic variables**; the elements of x, **non-basic variables**. The additive term d in the objective function clearly does not affect the optimality of a solution $(\hat{x}, \hat{y}) \in \mathbb{R}^n \times \mathbb{R}^m$—any solution $(\hat{x}, \hat{y})$ is optimal for $d = d_1$ if and only if it is optimal for $d = d_2$—and is normally not included in the definition of standard form. The reason we include it is because it allows the Simplex Algorithm to explicitly keep track of the objective function value of the current solution: The Simplex Algorithm constructs a sequence of equivalent LPs in standard form and corresponding feasible solutions $(\hat{x}, \hat{y})$ with the property that $\hat{x} = 0$. Since the variables in $\hat{y}$ do not appear in the objective function at all, this means that the solution $(\hat{x}, \hat{y})$ has objective function value d.

Proposition 3. *Every linear program in canonical form can be transformed into an equivalent linear program in standard form.*

Proof. Consider the LP (1) in canonical form and the LP (3) in standard form with $d = 0$. If $\hat{x}$ is a feasible solution of (1), it can be extended to a feasible solution of (3) by setting $\hat{y}_i = b_i - \sum_{j=1}^{n} a_{ij} \hat{x}_j$. This solution satisfies $\hat{y} \ge 0$ because $b_i \ge \sum_{j=1}^{n} a_{ij} \hat{x}_j$, for any feasible solution $\hat{x}$ of (1). Thus, $(\hat{x}, \hat{y})$ is a feasible solution of (3), and it is easy to see that it achieves the same objective function value as $\hat{x}$

viewed as a solution of (1).

If $(\hat{x}, \hat{y})$ is a feasible solution of (3), then $\hat{x}$ is a feasible solution of (1) because, for all $i \in [m]$, $\hat{y}_i = b_i - \sum_{j=1}^{n} a_{ij}\hat{x}_j$ and $\hat{y}_i \geq 0$, so $b_i \geq \sum_{j=1}^{n} a_{ij}\hat{x}_j$. Moreover, $\hat{x}$ as a solution of (1) achieves the same objective function value as $(\hat{x}, \hat{y})$ as a solution of (3).

This shows that (1) and (3) are equivalent LPs. $\square$

Example 4. To continue Ex. 2, we can convert the LP (4) into an equivalent LP in standard form by introducing a slack variable $y_{u,v}$ for every edge $(u,v) \in E$ and rewriting (4) as

$$
\begin{aligned}
&\text{Maximize } \sum_{u \in V} b_u p'_u - \sum_{u \in V} b_u p''_u - \sum_{(u,v) \in E_b} c_{u,v} s_{u,v} \\
&\text{s.t. } p'_u - p''_u - p'_v + p''_v - s_{u,v} + y_{u,v} = q_{u,v} && \forall (u,v) \in E_b \\
&\qquad p'_u - p''_u - p'_v + p''_v = y_{u,v} = -q_{u,v} && \forall (u,v) \in E_u \\
&\qquad p'_u \geq 0 && \forall u \in V \\
&\qquad p''_u \geq 0 && \forall u \in V \\
&\qquad s_{u,v} \geq 0 && \forall (u,v) \in E_b \\
&\qquad y_{u,v} \geq 0 && \forall (u,v) \in E_b \cup E_u
\end{aligned}
\tag{4}
$$

The standard form (3) can be rewritten as

$$
\begin{aligned}
&\text{Maximize } 0y + cx + d \\
&\text{s.t. } b = Iy + Ax \\
&\qquad x \geq 0 \\
&\qquad y \geq 0,
\end{aligned}
\tag{5}
$$

where I is the $m \times m$ identity matrix and 0 is the m-element 0-vector.

Now let c' be the concatenation of 0 and c,

$$
c' = (\underbrace{0, 0, \ldots, 0}_{m \text{ times}}, c_1, c_2, \ldots, c_n),
$$

let

$$
A' = [I|A] = \left(\begin{array}{cccc|cccc}
1 & 0 & \cdots & 0 & a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\
0 & 1 & \cdots & 0 & a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 1 & a_{m,1} & a_{m,2} & \cdots & a_{m,n}
\end{array}\right),
$$

and let

$$z = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_{m+n} \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \\ x_1 \\ x_2 \\ \vdots x_n \end{pmatrix}.$$

Then (5) becomes

$$\begin{aligned} \text{Maximize } & c'z + d \\ \text{s.t. } & b = A'z \\ & z \geq 0. \end{aligned} \tag{6}$$

An LP written in this form is said to be in standard form if it is equipped with a **basis** $B = \{j_1, \ldots, j_m\} \subseteq [m+n]$[3] of m indices such that for all $i \in [m]$, $c_{j_i} = 0$, $a_{ij_i} = 1$, and $a_{ij} = 0$, for all $j \in [m+n] \setminus \{j_i\}$. The variables $z_{j_1}, \ldots, z_{j_m}$ are the **basic variables** of the LP. To clearly associate each index $j_i \in B$ with the ith equality constraint in (6), we write B as a vector $(j_1, \ldots, j_m)$, not as a set, but we still use standard set notation such as $h \in B$ to indicate that an index h occurs in this vector. The above translation of (3) into (6) together with the basis $B = (1, \ldots, m)$ clearly satisfies these conditions.

References

Papadimitriou, Christos H. and Kenneth Steiglitz (1982). *Combinatorial Optimization: Algorithms and Complexity*. Prentice Hall.

[3] We may also refer to the set $\{z_{j_1}, \ldots, z_{j_m}\}$ as the basis. Whether we mean the basic variables or their indices will be clear from the context.