Chapter 3: Linear Programming

Ernest K. Ryu

MATH 164: Optimization
University of California, Los Angeles
Department of Mathematics

Last edited: March 10, 2025

Linear programming

A *linear program* (LP) is an optimization problem in which the objective function, equality constraints, and inequality constraints are all affine.

We can solve LPs (to globally optimality) efficiently.

- In complexity theory language, LPs are solvable in (weakly) polynomial time.
- LPs are convex optimization problems, i.e., (LP) $\subset$ (Cvx. Opt.).

Commonly used algorithms include interior point methods, first-order splitting methods, and the simplex method.

In this class, we will learn the simplex method.

(One should not conflate the problem with the algorithm used to solve it. LP is the mathematical problem, and the simplex algorithm is one of the several solution methods.)

Outline

Example: Advertising budget optimization

You have a budget of 10,000 dollars for advertising, and you want to split this amount among four channels. Assume you know the ROIs, the ratio of output (revenue gained) to input (ad spending) for each channel.

1. Search engine ads (e.g., Google). ROI: 25
2. Website/app displays (banner ads). ROI: 16
3. Online video ads (e.g., YouTube). ROI: 8
4. Pushed text ads (text messages). ROI: 6

Because different channels reach different audiences, your marketing guidelines for long-term growth require:

- A 2 and 3 combined must be at least 50% of the total budget,
- B 3 alone cannot exceed 30% of the total budget,
- C Minimum spending on 1 is 3,000, and
- D Minimum spending on 4 is 2,000.

We want to maximize total ROI.

Also, all ad buys cannot be negative.

Example: Advertising budget optimization

We can model this problem as a linear program

$$\begin{array}{ll} \underset{x_1,x_2,x_3,x_4\in\mathbb{R}}{\text{maximize}} & 25x_1 + 16x_2 + 8x_3 + 6x_4 \\ \text{subject to} & x_1 + x_2 + x_3 + x_4 \le 10000 \\ & 5000 \le x_2 + x_3 \\ & x_1 \ge 3000,\ x_2 \ge 0,\ 0 \le x_3 \le 3000,\ x_4 \ge 2000, \end{array}$$

where the decision variables x_1, x_2, x_3, x_4 represent the amounts spent on 1 search engines, 2 displays, 3 online videos, and 4 pushed text ads.

Of course, this is equivalent to the minimization problem

$$\begin{array}{ll} \underset{x_1,x_2,x_3,x_4\in\mathbb{R}}{\text{minimize}} & -25x_1 - 16x_2 - 8x_3 - 6x_4 \\ \text{subject to} & x_1 + x_2 + x_3 + x_4 \le 10000 \\ & 5000 \le x_2 + x_3 \\ & 3000 \le x_1,\ 0 \le x_2,\ 0 \le x_3 \le 3000,\ 2000 \le x_4. \end{array}$$

(When we talk about LP duality, we will see that it is convenient to adopt minimization, rather than maximization, as the standard convention.)

Example: Advertising budget optimization

$$
\begin{array}{ll}
\underset{x_1,x_2,x_3,x_4\in\mathbb{R}}{\text{maximize}} & 25x_1 + 16x_2 + 8x_3 + 6x_4 \\
\text{subject to} & x_1 + x_2 + x_3 + x_4 \le 10000 \\
& 5000 \le x_2 + x_3 \\
& 3000 \le x_1,\ 0 \le x_2,\ 0 \le x_3 \le 3000,\ 2000 \le x_4.
\end{array}
$$

Modern LP solvers, both commercial and open-source, are readily available, efficient, and robust. Using a solver, we obtain the solution

$$x_\star = (3000, 5000, 0, 2000).$$

Aside: Programming is planning

The term "programming" in linear *programming* doesn't refer to writing computer code. Instead, it comes from an older usage of the word meaning "to plan" or "to schedule."

(At a classical music concert, a "program" and is a booklet containing the plan for the concert.)

During World War II, linear programming was used to devise optimal plans for resource allocation, production schedules, or military logistics. It was about formulating a "program" (or plan) that would achieve the best possible outcome given a set of constraints.

(A computer "program" is a set of instructions (plans) for human computers or electronic computers to execute.)

Similarly, *mathematical programming* means (mathematical) optimization.

Example: Chebyshev approximation problem, or minimax approximation problem

Consider

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad \|Ax - b\|_\infty,$$

where $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. Assume $m > n$. In this case, we do not expect $Ax = b$ to be attainable. Goal is to minimize maximum deviation from $Ax = b$.

The original problem, as stated, is not an LP. But it can be transformed into (it is equivalent to) the following LP:

$$\begin{array}{ll} \underset{x \in \mathbb{R}^n,\, t \in \mathbb{R}}{\text{minimize}} & t \\ \text{subject to} & -t\mathbf{1} \leq Ax - b \leq t\mathbf{1}, \end{array}$$

where $\mathbf{1} \in \mathbb{R}^m$ is the vector of all 1's and $\leq$ to denote element-wise inequality.

We say two optimization problems are *equivalent* if we can easily obtain the solution from one problem with the solution from the other problem.

Example: Chebyshev approximation problem, or minimax approximation problem

Let's do the transformation step by step.

First, we show that

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad \|Ax - b\|_\infty \tag{P1}$$

is equivalent to

$$\begin{array}{ll} \underset{x \in \mathbb{R}^n,\, t \in \mathbb{R}}{\text{minimize}} & t \\ \text{subject to} & \|Ax - b\|_\infty \le t. \end{array} \tag{P2}$$

Let $p_\star^{\text{(P1)}}$ and $p_\star^{\text{(P2)}}$ be the optimal values for (P1) and (P2).

For any $x \in \mathbb{R}^n$, x is feasible for (P1) and (x, t) with $t = \|Ax - b\|_\infty$ is feasible for (P2). The two feasible points attain the same objective value. So any objective value (P1) can attain, (P2) can also attain it, and we conclude $p_\star^{\text{(P2)}} \le p_\star^{\text{(P1)}}$.

Example: Chebyshev approximation problem, or minimax approximation problem

We continue to show that

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad \|Ax - b\|_\infty \tag{P1}$$

is equivalent to

$$\begin{array}{ll} \underset{x \in \mathbb{R}^n,\, t \in \mathbb{R}}{\text{minimize}} & t \\ \text{subject to} & \|Ax - b\|_\infty \le t. \end{array} \tag{P2}$$

On the other hand, if (x, t) is feasible and attains objective value t for (P2), then x attains the objective value $\|Ax - b\|_\infty \le t$ for (P1). So any objective value (P2) can attain, (P1) can attain the same or better objective value, and we conclude $p_\star^{\text{(P1)}} \le p_\star^{\text{(P2)}}$.

So the two problems attain the same objective value $p_\star = p_\star^{\text{(P1)}} = p_\star^{\text{(P2)}}$.

Example: Chebyshev approximation problem, or minimax approximation problem

We continue to show that

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad \|Ax - b\|_\infty \tag{P1}$$

is equivalent to

$$\begin{array}{ll} \underset{x \in \mathbb{R}^n,\, t \in \mathbb{R}}{\text{minimize}} & t \\ \text{subject to} & \|Ax - b\|_\infty \le t. \end{array} \tag{P2}$$

If $x_\star$ is optimal (P1), then $(x_\star, t_\star)$ with $t_\star = \|Ax_\star - b\|_\infty$ attains the objective value $p_\star = \|Ax_\star - b\|_\infty$ and is therefore optimal for (P2).

If $(x_\star, t_\star)$ is optimal for (P2), then $p_\star = t_\star = \|Ax_\star - b\|_\infty$ (it cannot be that $t_\star > \|Ax_\star - b\|_\infty$), since otherwise we can improve the objective value. So, $x_\star$ attains objective value $p_\star = \|Ax_\star - b\|_\infty$ for (P1) is therefore optimal for (P1).

Example: Chebyshev approximation problem, or minimax approximation problem

In conclusion, if you solve

$$\underset{x\in\mathbb{R}^n}{\text{minimize}} \quad \|Ax-b\|_\infty \tag{P1}$$

and get a solution $x_\star$, then you can immediately compute $t_\star = \|Ax_\star - b\|_\infty$, and return $(x_\star, t_\star)$ as the solution to (P2).

Conversely, if you solve

$$\begin{array}{ll} \underset{x\in\mathbb{R}^n,\, t\in\mathbb{R}}{\text{minimize}} & t \\ \text{subject to} & \|Ax-b\|_\infty \le t \end{array} \tag{P2}$$

and get a solution $(x_\star, t_\star)$, then we can return $x_\star$ (discarding $t_\star$) as the solution to (P1).

(Often, the equivalence of optimization problems is argued informally because a formal/rigorous argument can become quite tedious, as is the case here. However, presenting a formal proof along with an explicit algorithm that transforms a solution of one problem into a solution helps to ensure correctness.)

Example: Chebyshev approximation problem, or minimax approximation problem

Finally, we argue that

$$\begin{array}{ll} \underset{x\in\mathbb{R}^n,\, t\in\mathbb{R}}{\text{minimize}} & t \\ \text{subject to} & \|Ax-b\|_\infty \le t \end{array} \tag{P2}$$

is equivalent to

$$\begin{array}{ll} \underset{x\in\mathbb{R}^n,\, t\in\mathbb{R}}{\text{minimize}} & t \\ \text{subject to} & -t\mathbf{1} \le Ax-b \le t\mathbf{1}. \end{array} \tag{P3}$$

This is because the constraint sets are, by definition, equal sets:

$$\{x \in \mathbb{R}^n,\, t \in \mathbb{R} : \|Ax-b\|_\infty \le t\} = \{x \in \mathbb{R}^n,\, t \in \mathbb{R} : -t\mathbf{1} \le Ax-b \le t\mathbf{1}\}.$$

Outline

LP applications

LP theory

Weak duality

Strong duality

Examples

Simplex method in tableau form

Standard form

The *standard form* of an LP has the form

$$\begin{array}{ll} \underset{x \in \mathbb{R}^n}{\text{minimize}} & c^\intercal x \\ \text{subject to} & Ax = b \\ & x \geq 0, \end{array}$$

where $x \in \mathbb{R}^n$ is the optimization variable and $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, and $c \in \mathbb{R}^n$ are problem data.

We use $\geq$ and $\leq$ to denote element-wise inequality of vectors, i.e. $x \geq 0$ means $x_i \geq 0$ for all $i = 1, \ldots, n$. (Just as $=$ between two vectors is interpreted element-wise.)

Many standard references on LPs and the simplex method use the standard form for simplicity. Indeed, all LPs can be converted to the standard form.

However, many practical problems are more convenient and natural to express in non-standard LP form. Also, it may be algorithmically inefficient to convert a given LP into the standard form.

Extended form

The *extended form* of an LP has the form:

$$\begin{array}{ll} \underset{x \in \mathbb{R}^n}{\text{minimize}} & c^\mathsf{T} x \\ \text{subject to} & Ax = b \\ & Cx \leq d \\ & \ell \leq x. \end{array}$$

where $A \in \mathbb{R}^{m \times n}$, $C \in \mathbb{R}^{p \times n}$, $b \in \mathbb{R}^m$, $d \in \mathbb{R}^p$, and $\ell \in \mathbb{R}^n$.
The extended form also allows one to specify linear inequality constraints $Cx \leq d$ and more flexible lower bounds $\ell \leq x$.

The flexibility of the extended form makes it more convenient.
Mathematically speaking, however, the extended form is not more general since an LP in extended form can be converted into standard form.

Transformation into standard form

We shall convert the extended form LP

$$\begin{array}{ll}\underset{x\in\mathbb{R}^n}{\text{minimize}} & c^\intercal x \\ \text{subject to} & Ax = b \\ & Cx \le d \\ & \ell \le x\end{array}$$

into standard form. First, perform the change of variables $y = x - \ell$:

$$\begin{array}{ll}\underset{y\in\mathbb{R}^n}{\text{minimize}} & c^\intercal y + c^\intercal \ell \\ \text{subject to} & Ay = \tilde{b} \\ & Cy \le \tilde{d} \\ & y \ge 0\end{array}$$

where $\tilde{b} = b - A\ell$ and $\tilde{d} = d - C\ell$. Note that $c^\intercal \ell$ is a constant.

Transformation into standard form

Next, we argue that

$$\begin{array}{ll} \underset{y\in\mathbb{R}^n}{\text{minimize}} & c^\mathsf{T} y + c^\mathsf{T}\ell \\ \text{subject to} & Ay = \tilde{b} \\ & Cy - \tilde{d} \le 0 \\ & y \ge 0. \end{array}$$

is equivalent to

$$\begin{array}{ll} \underset{y\in\mathbb{R}^n,\, s\in\mathbb{R}^p}{\text{minimize}} & c^\mathsf{T} y + c^\mathsf{T}\ell \\ \text{subject to} & Ay = \tilde{b} \\ & Cy - \tilde{d} = -s \\ & s \ge 0,\, y \ge 0. \end{array}$$

The trick is referred to as introducing a *slack variable* s.

A downside of introducing a slack variable is that the problem dimension increases, and this can make the algorithm less efficient.

Transformation into standard form

Finally

$$\begin{array}{ll} \underset{y\in\mathbb{R}^n,\, s\in\mathbb{R}^p}{\text{minimize}} & c^\intercal y + c^\intercal \ell \\ \text{subject to} & Ay = \tilde{b} \\ & Cy - \tilde{d} = -s \\ & s \geq 0,\ y \geq 0 \end{array}$$

is equivalent to

$$\begin{array}{ll} \underset{(y,s)\in\mathbb{R}^{n+p}}{\text{minimize}} & \begin{bmatrix} c^\intercal \\ 0_p \end{bmatrix}^\intercal \begin{bmatrix} y \\ s \end{bmatrix} \\ \text{subject to} & \begin{bmatrix} A & 0 \\ C & I_{p\times p} \end{bmatrix} \begin{bmatrix} y \\ s \end{bmatrix} = \begin{bmatrix} \tilde{b} \\ \tilde{d} \end{bmatrix} \\ & \begin{bmatrix} y \\ s \end{bmatrix} \geq 0, \end{array}$$

where $0_p \in \mathbb{R}^p$ is the vector of all 0's and $I_{p\times p} \in \mathbb{R}^{p\times p}$ is the $p \times p$ identity matrix and we removed the constant from the objective function since it does not affect the solution (but it does affect the optimal value by that constant amount). We are now in standard form.

General form

The *general form* offers further flexibility in specifying lower and upper limits on both Ax and x itself:

$$\begin{array}{ll} \underset{x \in \mathbb{R}^n}{\text{minimize}} & c^\intercal x \\ \text{subject to} & L \leq Ax \leq U \\ & \ell \leq x \leq u, \end{array}$$

where $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$.

We let L and U be length m "vectors" satisfying $L \leq U$, but we allow $L_i = -\infty$ or $U_i = +\infty$ for any $i = 1, \ldots, m$ to indicate no constraint in that direction. So $-\infty \leq a_i^\intercal x \leq U_i$ means $a_i^\intercal x \leq U_i$, and $L_i \leq a_i^\intercal x \leq \infty$ means $L_i \leq a_i^\intercal x$. Likewise, we let ℓ and u be length n "vectors" satisfying $\ell \leq u$ that can take $-\infty$ and $+\infty$ values. (To clarify, no bound in the standard and extended forms is allowed to take $\pm\infty$ values.)

Equality constraints are encoded by setting $-\infty < L_i = U_i < \infty$ or $-\infty < \ell_i = u_i < \infty$.

Transformation into extended form: Case 1

As before, we can transform a general form LP into the extended form, and this can, in turn, be transformed into the standard form.

For the sake of simplicity, assume $-\infty < L < U < \infty$ and $-\infty < \ell < u < \infty$. Then,

$$\begin{array}{ll} \underset{x \in \mathbb{R}^n}{\text{minimize}} & c^\intercal x \\ \text{subject to} & L \le Ax \le U \\ & \ell \le x \le u \end{array}$$

is equivalent to

$$\begin{array}{ll} \underset{x, x' \in \mathbb{R}^n}{\text{minimize}} & c^\intercal x \\ \text{subject to} & x + x' = 0 \\ & \begin{bmatrix} A \\ -A \end{bmatrix} x \le \begin{bmatrix} U \\ -L \end{bmatrix} \\ & \begin{bmatrix} \ell \\ -u \end{bmatrix} \le \begin{bmatrix} x \\ x' \end{bmatrix}. \end{array}$$

Transformation into extended form: Case 1

Further,

$$\begin{array}{ll} \underset{x, x' \in \mathbb{R}^n}{\text{minimize}} & c^\mathsf{T} x \\ \text{subject to} & x + x' = 0 \\ & \begin{bmatrix} A \\ -A \end{bmatrix} x \leq \begin{bmatrix} U \\ -L \end{bmatrix} \\ & \begin{bmatrix} \ell \\ -u \end{bmatrix} \leq \begin{bmatrix} x \\ x' \end{bmatrix} \end{array}$$

is equivalent to

$$\begin{array}{ll} \underset{x, x' \in \mathbb{R}^n}{\text{minimize}} & \begin{bmatrix} c \\ 0 \end{bmatrix}^\mathsf{T} \begin{bmatrix} x \\ x' \end{bmatrix} \\ \text{subject to} & \begin{bmatrix} I & I \end{bmatrix} \begin{bmatrix} x \\ x' \end{bmatrix} = 0 \\ & \begin{bmatrix} A & 0 \\ -A & 0 \end{bmatrix} \begin{bmatrix} x \\ x' \end{bmatrix} \leq \begin{bmatrix} U \\ -L \end{bmatrix} \\ & \begin{bmatrix} \ell \\ -u \end{bmatrix} \leq \begin{bmatrix} x \\ x' \end{bmatrix}. \end{array}$$

We are now in extended form.

Transformation into extended form: Case 2

Consider another case where $-\infty < L < U < \infty$, $\ell_i = -\infty$ for $i = 1, \ldots, n$, and $u_i = +\infty$ for $i = 1, \ldots, n$. Then,

$$\begin{array}{ll} \underset{x \in \mathbb{R}^n}{\text{minimize}} & c^\mathsf{T} x \\ \text{subject to} & L \le Ax \le U \\ & \ell \le x \le u \end{array}$$

is equivalent to

$$\begin{array}{ll} \underset{x \in \mathbb{R}^n}{\text{minimize}} & c^\mathsf{T} x \\ \text{subject to} & L \le Ax \le U. \end{array}$$

Transformation into extended form: Case 2

Note that x has no direct upper or lower bounds. We deal with this by splitting x into the positive and negative parts, i.e., $x = x_+ - x_-$.

Specifically,

$$\begin{array}{ll} \underset{x \in \mathbb{R}^n}{\text{minimize}} & c^\intercal x \\ \text{subject to} & L \leq Ax \leq U. \end{array}$$

is equivalent to

$$\begin{array}{ll} \underset{x_+, x_- \in \mathbb{R}^n}{\text{minimize}} & c^\intercal (x_+ - x_-) \\ \text{subject to} & L \leq Ax_+ - Ax_- \leq U \\ & 0 \leq x_+,\, 0 \leq x_-. \end{array}$$

Transformation into extended form: Case 2

Further,

$$\begin{array}{ll} \underset{x_+, x_- \in \mathbb{R}^n}{\text{minimize}} & c^\mathsf{T}(x_+ - x_-) \\ \text{subject to} & L \leq Ax_+ - Ax_- \leq U \\ & 0 \leq x_+,\, 0 \leq x_- \end{array}$$

is equivalent to

$$\begin{array}{ll} \underset{x_+, x_- \in \mathbb{R}^n}{\text{minimize}} & \begin{bmatrix} c \\ -c \end{bmatrix}^\mathsf{T} \begin{bmatrix} x_+ \\ x_- \end{bmatrix} \\ \text{subject to} & \begin{bmatrix} A & -A \\ -A & A \end{bmatrix} \begin{bmatrix} x_+ \\ x_- \end{bmatrix} \leq \begin{bmatrix} U \\ -L \end{bmatrix} \\ & 0 \leq \begin{bmatrix} x_+ \\ x_- \end{bmatrix}. \end{array}$$

We are now in extended form.

The transformation of a general general form LP into extended form can be done by combining the techniques of the demonstrated Cases 1 and 2.

Convexity

LPs have the following convexity properties.

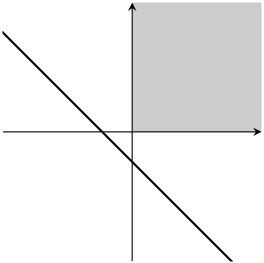
- The objective function $c^\mathsf{T} x$ is convex.
- The feasible set is convex, i.e., if x_1 and x_2 are feasible, then $\theta x_1 + (1-\theta)x_2$ is feasible for $\theta \in [0,1]$.
- The optimal solution set is convex, i.e., if x_1 and x_2 are optimal, then $\theta x_1 + (1-\theta)x_2$ is optimal for $\theta \in [0,1]$.

We leave the proof as an exercise.

Infeasible problems

We say an LP is *infeasible* if it has no feasible point. For example, the standard form LP

$$\begin{array}{ll} \underset{(x,y)\in\mathbb{R}^2}{\text{minimize}} & \cdots \\ \text{subject to} & \begin{bmatrix}1 & 1\end{bmatrix}\begin{bmatrix}x \\ y\end{bmatrix} = -1 \\ & x \geq 0,\ y \geq 0. \end{array}$$



is infeasible.

If the problem is infeasible, we write $p^\star = \infty$ for the optimal value.

People specify incompatible constraints all the time, so we shall consider infeasible instances as a legitimate possibility within the LP framework.

Unbounded problems

Consider

$$\begin{array}{ll} \underset{x\in\mathbb{R}^n}{\text{minimize}} & c^\intercal x \\ \text{subject to} & Ax = b \\ & x \geq 0, \end{array}$$

and assume the problem is feasible with feasible point x_0. (So, $p_\star < \infty$.)

Further assume there is a direction $v \in \mathbb{R}^n$ such that $Av = 0$, $v \geq 0$, and $c^\intercal v < 0$. Then, $x_0 + \alpha v$ for $\alpha > 0$ is feasible and has objective value

$$c^\intercal x_0 + \alpha c^\intercal v \to -\infty \qquad \text{as } \alpha \to \infty.$$

So, $p_\star = -\infty$, and we say the problem is *unbounded*. Such a $v \in \mathbb{R}^n$ is called a *direction of unboundedness*.

(Using duality, we will see that the converse is true: if $p_\star = -\infty$, then the LP is feasible and there is a direction of unboundedness.)

Unbounded problems

As an aside, because LPs have linear objectives, the optimization problem is meaningful only with constraints.

Consider an unconstrained LP,

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad c^\mathsf{T} x.$$

If $c \neq 0$, then $v = c$ would be a direction of unboundedness and $p_\star = -\infty$. If $c = 0$, then the problem is even less interesting.

Outline

Dual LP

Consider the standard form LP

$$\begin{array}{ll} \underset{x \in \mathbb{R}^n}{\text{minimize}} & c^\intercal x \\ \text{subject to} & Ax = b \\ & x \geq 0, \end{array} \tag{P}$$

where $x \in \mathbb{R}^n$ is the optimization variable and $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, and $c \in \mathbb{R}^n$ are problem data. We shall call this the *primal problem* and write the optimal value as $p_\star \in [-\infty, \infty]$.

Consider

$$\begin{array}{ll} \underset{y \in \mathbb{R}^m}{\text{maximize}} & b^\intercal y \\ \text{subject to} & A^\intercal y \leq c. \end{array} \tag{D}$$

We shall call this the *dual problem*, and write the optimal value as $d_\star \in [-\infty, \infty]$

Weak duality for standard form

$$\begin{array}{ll} \underset{x\in\mathbb{R}^n}{\text{minimize}} & c^\intercal x \\ \text{subject to} & Ax = b \\ & x \geq 0 \end{array} \quad \overset{\text{dual}}{\longleftrightarrow} \quad \begin{array}{ll} \underset{y\in\mathbb{R}^m}{\text{maximize}} & b^\intercal y \\ \text{subject to} & A^\intercal y \leq c \end{array}$$

Theorem (Weak duality).

The optimal values of the primal and dual problems satisfy

$$d^\star \leq p^\star.$$

Proof. If $d_\star = -\infty$ or $p_\star = \infty$, there is nothing to show. So assume $-\infty < d_\star$ and $p^\star < \infty$, i.e., the dual and primal problems are feasible. Let x and y be primal and dual feasible points. Then,

$$(y^\intercal A - c^\intercal)x = (-)(+) \leq 0$$

where $(-)$ and $(+)$ means the vectors are element-wise non-positive and non-negative. Finally, we conclude

$$b^\intercal y = y^\intercal Ax \leq c^\intercal x.$$

$\square$

Consequences of weak duality

Corollary.

Consider the primal-dual correspondence

$$\begin{array}{ll} \underset{x \in \mathbb{R}^n}{\text{minimize}} & c^\mathsf{T} x \\ \text{subject to} & Ax = b \\ & x \geq 0 \end{array} \quad \overset{\text{dual}}{\longleftrightarrow} \quad \begin{array}{ll} \underset{y \in \mathbb{R}^m}{\text{maximize}} & b^\mathsf{T} y \\ \text{subject to} & A^\mathsf{T} y \leq c \end{array}$$

1. *If the primal problem is feasible but unbounded* $p_\star = -\infty$, *then the dual problem is infeasible.*
2. *If the dual problem is feasible but unbounded* $d_\star = +\infty$, *then the primal problem is infeasible.*
3. *If* (x, y) *are feasible and* $b^\mathsf{T} y = c^\mathsf{T} x$, *then both are optimal and* $d_\star = b^\mathsf{T} y = c^\mathsf{T} x = p_\star$.

Certificate of optimality

$$\begin{array}{ll} \underset{x \in \mathbb{R}^n}{\text{minimize}} & c^\intercal x \\ \text{subject to} & Ax = b \\ & x \geq 0 \end{array} \quad \overset{\text{dual}}{\longleftrightarrow} \quad \begin{array}{ll} \underset{y \in \mathbb{R}^m}{\text{maximize}} & b^\intercal y \\ \text{subject to} & A^\intercal y \leq c \end{array}$$

3. If (x, y) are feasible and $b^\intercal y = c^\intercal x$, then both are optimal and $d_\star = b^\intercal y = c^\intercal x = p_\star$.

Point #3 is very useful because it provides a *certificate* of optimality. Otherwise, if I assert that an x is optimal, how would you trust me?

In unconstrained differentiable convex minimization, if I say $x_\star$ minimizes f, you can check it by seeing that $\nabla f(x_\star) = 0$.

But, is this ever going to happen? We've shown $d_\star \leq p_\star$, but perhaps $d_\star < p_\star$ is the norm? (Spoiler, $d_\star = p_\star$ usually holds for LPs.)

Weak duality for extended form

Similar primal-dual correspondence for the extended form:

$$\begin{array}{ll} \underset{x\in\mathbb{R}^n}{\text{minimize}} & c^\mathsf{T} x \\ \text{subject to} & Ax = b \\ & Cx \le d \\ & \ell \le x \end{array} \quad \overset{\text{dual}}{\longleftrightarrow} \quad \begin{array}{ll} \underset{y_b\in\mathbb{R}^m,\, y_d\in\mathbb{R}^p,\, y_\ell\in\mathbb{R}^x}{\text{maximize}} & b^\mathsf{T} y_b + d^\mathsf{T} y_d + \ell^\mathsf{T} y_\ell \\ \text{subject to} & A^\mathsf{T} y_b + C^\mathsf{T} y_d + y_\ell = c \\ & y_d \le 0,\ y_\ell \ge 0, \end{array}$$

where $A \in \mathbb{R}^{m\times n}$, $C \in \mathbb{R}^{p\times n}$, $b \in \mathbb{R}^m$, $d \in \mathbb{R}^p$, and $\ell \in \mathbb{R}^n$.

Theorem (Weak duality).

The optimal values of the primal and dual problems satisfy

$$d^\star \le p^\star.$$

Proof. Exercise. $\square$

Weak duality for general form

Similar primal-dual correspondence for the extended form:

$$\begin{array}{ll} \underset{x\in\mathbb{R}^n}{\text{minimize}} & c^\intercal x \\ \text{subject to} & L \le Ax \le U \\ & \ell \le x \le u \end{array} \tag{P}$$

and

$$\begin{array}{ll} \underset{\substack{y_L, y_U \in\mathbb{R}^m \\ y_\ell, y_u \in\mathbb{R}^n}}{\text{maximize}} & L^\intercal y_L - U^\intercal y_U + \ell^\intercal y_\ell - u^\intercal y_u \\ \text{subject to} & A^\intercal y_L - A^\intercal y_U + y_\ell - y_u = c \\ & y_L, y_U, y_\ell, y_u \ge 0, \end{array} \tag{D}$$

where $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$.

For the dual problem, use the convention $0 \cdot (-\infty) = 0 \cdot \infty = 0$, $\alpha \cdot \pm\infty = \pm\infty$ for $\alpha \neq 0$, where the $\pm$ signs follow the obvious convension. This implies that the y-value must be 0 for all infinite L, U, ℓ, u values, since otherwise the objective function would be $-\infty$, the most undesirable value.

Weak duality for general form

$$\begin{array}{ll} \underset{x \in \mathbb{R}^n}{\text{minimize}} & c^\intercal x \\ \text{subject to} & L \leq Ax \leq U \\ & \ell \leq x \leq u \end{array} \tag{P}$$

$$\begin{array}{ll} \underset{\substack{y_L, y_U \in \mathbb{R}^m \\ y_\ell, y_u \in \mathbb{R}^n}}{\text{maximize}} & L^\intercal y_L - U^\intercal y_U + \ell^\intercal y_\ell - u^\intercal y_u \\ \text{subject to} & A^\intercal y_L - A^\intercal y_U + y_\ell - y_u = c \\ & y_L, y_U, y_\ell, y_u \geq 0, \end{array} \tag{D}$$

Theorem (Weak duality).

The optimal values of the primal and dual problems satisfy

$$d^\star \leq p^\star.$$

Proof. Exercise. □

Maximin-minimax derivation of dual

We introduced dual LPs corresponding to the primal LPs out of nowhere.

Once the primal and dual problems are stated, it is not too difficult to show weak duality. But, where does the dual problem come from?

Answer) We can derive the dual using the maximin-minimax inequality and a well-chosen Lagrangian.

Lemma (Maximin-minimax inequality).

Let $L\colon X \times Y \to \mathbb{R}$ be an arbitrary function. Then,

$$\sup_{y \in Y} \inf_{x \in X} L(x, y) \le \inf_{x \in X} \sup_{y \in Y} L(x, y).$$

Proof. This follows from

$$\begin{aligned} L(x, y) &\le \sup_{y \in Y} L(x, y), \qquad \forall x \in X,\ y \in Y \\ \inf_{x \in X} L(x, y) &\le \inf_{x \in X} \sup_{y \in Y} L(x, y), \qquad \forall y \in Y \\ \sup_{y \in Y} \inf_{x \in X} L(x, y) &\le \inf_{x \in X} \sup_{y \in Y} L(x, y). \end{aligned}$$

$\square$

General weak duality

Let $L \colon X \times Y \to \mathbb{R}$ be an arbitrary function. Define $f \colon X \to \mathbb{R} \cup \{\infty\}$ and $g \colon Y \to \mathbb{R} \cup \{-\infty\}$ as

$$f(x) = \sup_{y \in Y} L(x, y) \qquad g(y) = \inf_{x \in X} L(x, y)$$

We call

$$\underset{x \in X}{\text{minimize}} \quad f(x) \tag{P}$$

the primal problem with optimal value $p_\star \in [-\infty, \infty]$

$$\underset{y \in Y}{\text{maximize}} \quad g(y) \tag{D}$$

the dual problem with optimal value $d_\star \in [-\infty, \infty]$.

Theorem (General weak duality).

For the primal and dual optimization problems defined above, we have

$$d_\star = \sup_{y \in Y} g(y) \leq \inf_{x \in X} f(x) = p_\star.$$

Proof. Immediate consequence of the maximin-minimax inequality. □

Primal-dual pair via Lagrangian L

$$f(x) = \sup_{y \in Y} L(x, y) \qquad \overset{\text{dual}}{\longleftrightarrow} \qquad g(y) = \inf_{x \in X} L(x, y)$$

$$\underset{x \in X}{\text{minimize}} \quad f(x) \qquad\qquad\qquad \underset{y \in Y}{\text{maximize}} \quad g(y)$$

We call L a *Lagrangian*. (Terminology comes from method of Lagrange multipliers.)

Pick any L, and we get a primal-dual pair of problems.

If we pick L such that the primal problem becomes our problem of interest, then we have a useful corresponding dual problem.

Maximizing linear functions over $\mathbb{R}^n$

We quickly establish two simple lemmas.

Lemma.

Let $v \in \mathbb{R}^n$. Then,

$$\inf_{x \in \mathbb{R}^n} v^\intercal x = \begin{cases} 0 & \textit{if } v = 0 \\ -\infty & \textit{otherwise.} \end{cases}$$

Proof. If $v = 0$, then $v^\intercal x = 0$ and the supremum is 0. If $v \neq 0$, then with $x = -\alpha v$, we have $v^\intercal x = -\alpha \|v\|^2 \to -\infty$ as $\alpha \to \infty$. $\square$

Maximizing linear functions over $\mathbb{R}^n_+$

Let

$$\mathbb{R}^n_+ = \{x \in \mathbb{R}^n \mid x \geq 0\}$$

be the n-dimensional *nonnegative orthant*.

Lemma.

Let $v \in \mathbb{R}^n$. Then,

$$\inf_{x \in \mathbb{R}^n_+} v^\intercal x = \begin{cases} 0 & \textit{if } v \in \mathbb{R}^n_+ \\ -\infty & \textit{otherwise.} \end{cases}$$

Proof. Note that we are minimizing over $x \geq 0$. If $v \geq 0$, then $v^\intercal x \geq 0$, so the infimum of 0 is attained at $x = 0$. If $v \ngeq 0$, then there is an index i such that $v_i < 0$. Setting $x = \alpha e_i$, where e_i is the i-th unit vector (all 0's except a 1 at the i-th coordinate), we have $v^\intercal x = \alpha v_i \to -\infty$ as $\alpha \to \infty$. $\square$

Deriving dual LP from Lagrangian

$$\begin{array}{ll} \underset{x \in \mathbb{R}^n}{\text{minimize}} & c^\mathsf{T} x \\ \text{subject to} & Ax = b \\ & x \geq 0 \end{array} \quad \overset{\text{dual}}{\longleftrightarrow} \quad \begin{array}{ll} \underset{y \in \mathbb{R}^m}{\text{maximize}} & b^\mathsf{T} y \\ \text{subject to} & A^\mathsf{T} y \leq c \end{array}$$

Let

$$\begin{aligned} L(x, y, s) &= c^\mathsf{T} x + y^\mathsf{T}(Ax - b) - s^\mathsf{T} x \\ &= (c - A^\mathsf{T} y - s)^\mathsf{T} x + b^\mathsf{T} y, \end{aligned}$$

where x is the primal variable and (y, s) are the dual variables. Then,

$$f(x) = \sup_{y \in \mathbb{R}^m,\, s \in \mathbb{R}^n_+} L(x, y, s) = \begin{cases} c^\mathsf{T} x & \text{if } Ax = b,\ x \geq 0 \\ +\infty & \text{otherwise} \end{cases}$$

and

$$g(y, s) = \inf_{x \in \mathbb{R}^n} L(x, y, s) = \begin{cases} b^\mathsf{T} y & \text{if } c - A^\mathsf{T} y - s = 0 \\ -\infty & \text{otherwise.} \end{cases}$$

Deriving dual LP from Lagrangian

$$f(x) = \sup_{y \in \mathbb{R}^m,\, s \in \mathbb{R}^n_+} L(x, y, s) = \begin{cases} c^\intercal x & \text{if } Ax = b,\ x \geq 0 \\ +\infty & \text{otherwise.} \end{cases}$$

We see that $\inf_{x \in \mathbb{R}^n} f(x)$ is equivalent to the primal problem

$$\begin{array}{ll} \underset{x \in \mathbb{R}^n}{\text{minimize}} & c^\intercal x \\ \text{subject to} & Ax = b \\ & x \geq 0. \end{array}$$

Our choice of L is useful in this context because
$f(x) = \sup_{y \in \mathbb{R}^m,\, s \in \mathbb{R}^n_+} L(x, y, s)$ recovers the primal LP.

Deriving dual LP from Lagrangian

$$g(y, s) = \inf_{x \in \mathbb{R}^n} L(x, y, s) = \begin{cases} b^\intercal y & \text{if } c - A^\intercal y - s = 0 \\ -\infty & \text{otherwise.} \end{cases}$$

We see that $\sup_{y \in \mathbb{R}^m,\, s \in \mathbb{R}^n_+} g(y, s)$ is equivalent to

$$\begin{array}{ll} \underset{y \in \mathbb{R}^m,\, s \in \mathbb{R}^n}{\text{maximize}} & b^\intercal y \\ \text{subject to} & c - A^\intercal y = s,\ s \geq 0, \end{array}$$

which is equivalent to the dual problem

$$\begin{array}{ll} \underset{y \in \mathbb{R}^m}{\text{maximize}} & b^\intercal y \\ \text{subject to} & A^\intercal y \leq c \end{array}$$

upon eliminating s. (So, this is a *derivation* of the dual LP.)

Finally, we conclude $d_\star \leq p_\star$.

Outline

Strong duality

Previously, we stated weak duality: $d_\star \leq p_\star$. In most cases, however, the inequality holds with equality.

Theorem (Informal).

Usually,

$$d_\star = p_\star$$

holds between the primal and dual LPs.

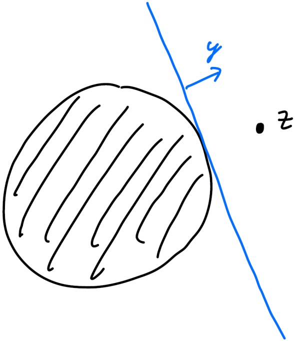
This is a very powerful result of linear programming and more broadly for (constrained) convex optimization.

Separating hyperplane theorem

Theorem (Separating hyperplane theorem).

Let $C \subset \mathbb{R}^n$ *be a nonempty closed convex set, and let* $z \in \mathbb{R}^n$*. If* $z \notin C$*, then there is a* $(y, \beta) \in \mathbb{R}^n \times \mathbb{R}$ *such that*

$$\begin{aligned} y^\mathsf{T} x &\le \beta, \qquad \forall\, x \in C \\ y^\mathsf{T} z &> \beta. \end{aligned}$$



Theorem (Separating hyperplane theorem).

Let $C \subset \mathbb{R}^n$ be a nonempty closed convex set, and let $z \in \mathbb{R}^n$. If $z \notin C$, then there is a $(y, \beta) \in \mathbb{R}^n \times \mathbb{R}$ such that

$$\begin{aligned} y^\mathsf{T} x &\leq \beta, \qquad \forall\, x \in C \\ y^\mathsf{T} z &> \beta. \end{aligned}$$

Proof. Let $\Pi(z)$ be the projection of z onto C, and let $y = z - \Pi(z)$. Note, $y \neq 0$, since $z \notin C$. By the projection theorem,

$$\langle x - \Pi(z), y \rangle \leq 0, \qquad \forall\, x \in C.$$

If we let $\beta = \langle \Pi(z), y \rangle$, then

$$y^\mathsf{T} x \leq \beta, \qquad \forall\, x \in C,$$

and

$$y^\mathsf{T} z = \langle z - \Pi(z), z \rangle = \underbrace{\langle z - \Pi(z), z - \Pi(z) \rangle}_{=\|y\|^2 > 0} + \underbrace{\langle z - \Pi(z), \Pi(z) \rangle}_{=\beta} > \beta. \qquad \square$$

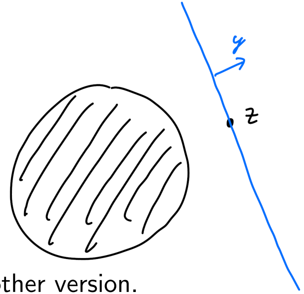
Separating hyperplane theorem

There are many variants of the separating hyperplane theorem.

Theorem (Separating hyperplane theorem).

Let $C \subset \mathbb{R}^n$ be a nonempty closed convex set, and let $z \in \mathbb{R}^n$. If $z \notin C$, then there is a $(y, \beta) \in \mathbb{R}^n \times \mathbb{R}$ such that

$$\begin{aligned} y^\mathsf{T} x &< \beta, \qquad \forall\, x \in C \\ y^\mathsf{T} z &\geq \beta. \end{aligned}$$



Proof. Similar to the other version. □

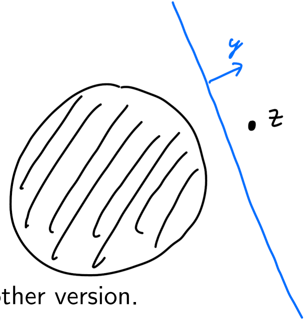
Separating hyperplane theorem

There are many variants of the separating hyperplane theorem.

Theorem (Separating hyperplane theorem).

Let $C \subset \mathbb{R}^n$ be a nonempty closed convex set, and let $z \in \mathbb{R}^n$. If $z \notin C$, then there is a $(y, \beta) \in \mathbb{R}^n \times \mathbb{R}$ such that

$$\begin{aligned} y^\mathsf{T} x &< \beta, \qquad \forall\, x \in C \\ y^\mathsf{T} z &> \beta. \end{aligned}$$



Proof. Similar to the other version. □

Farkas' lemma

Farkas' lemma is fundamental to establishing strong duality between LPs.

Lemma (Farkas' lemma).

Given $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$, exactly one of the following holds:

- *There exists $x \in \mathbb{R}^n$ such that $Ax = b$ and $x \geq 0$,*
- *There exists $y \in \mathbb{R}^m$ such that $A^\intercal y \leq 0$ and $b^\intercal y > 0$.*

(If one statement is false, the other must be true.)

Such a result is referred to as a *theorem of alternatives*, meaning it is a theorem stating that exactly one of two statements hold true.

Farkas' lemma

In computer programming and Boolean logic, the *exclusive or* operator written as XOR has the truth table

A	B	A (XOR) B
0	0	0
0	1	1
1	0	1
1	1	0

Farkas' lemma is often expressed with the XOR operator as follows,

Lemma (Farkas' lemma).

Let $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. Then,

- *there exists $x \in \mathbb{R}^n$ such that $Ax = b$ and $x \geq 0$*

  XOR

- *there exists $y \in \mathbb{R}^m$ such that $A^\intercal y \leq 0$ and $b^\intercal y > 0$.*

Alternatives as a certificate of infeasibility

Lemma (Farkas' lemma).

Let $A \in \mathbb{R}^{m \times n}$ *and* $b \in \mathbb{R}^m$. *Then,*

- *there exists* $x \in \mathbb{R}^n$ *such that* $Ax = b$ *and* $x \geq 0$

  XOR

- *there exists* $y \in \mathbb{R}^m$ *such that* $A^\intercal y \leq 0$ *and* $b^\intercal y > 0$.

If $[Ax = b$ and $x \geq 0]$ is infeasible, the y satisfying $[A^\intercal y \leq 0$ and $b^\intercal y > 0]$ provides a certificate (proof) of infeasibility.

Proof of Farkas

Lemma (Farkas' lemma).

Let $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. Then,

- *there exists $x \in \mathbb{R}^n$ such that $Ax = b$ and $x \geq 0$*

  XOR

- *there exists $y \in \mathbb{R}^m$ such that $A^\intercal y \leq 0$ and $b^\intercal y > 0$.*

Proof. There are 4 cases.

	$\exists x$	$\exists y$
Case 1	✗	✗
Case 2	✓	✗
Case 3	✗	✓
Case 4	✓	✓

In Case 4, there is a y satisfying $[A^\intercal y \leq 0$ and $b^\intercal y > 0]$ and an x such that $Ax = b$ and $x \geq 0$. Then, $0 < b^\intercal y = \underbrace{x^\intercal}_{\geq 0} \underbrace{A^\intercal y}_{\leq 0} \leq 0$ and we have a contradiction. So Case 4 cannot happen.

In Cases 2 and 3, we are happy.

It remains to show that Case 1 cannot happen.

Proof of Farkas

Assume there is no $x \in \mathbb{R}^n$ such that $Ax = b$ and $x \geq 0$. In other words, assume

$$b \notin S \stackrel{\text{def}}{=} \{Ax \,|\, x \geq 0\}.$$

Clearly, $0 \in S$ and it can be shown that S is closed and convex.

Since $b \notin S$, the separating hyperplane theorem tells us that

$$\left(\exists\, y \in \mathbb{R}^m,\, \beta \in \mathbb{R} \,:\, \begin{array}{l} y^\mathsf{T} v \leq \beta,\ \forall\, v \in S \\ y^\mathsf{T} b > \beta \end{array} \right)$$

Since $0 \in S$, we must have $\beta \geq 0$. So,

$$\left(\exists\, y \in \mathbb{R}^m,\, \beta \geq 0 \,:\, \begin{array}{l} y^\mathsf{T} v \leq \beta,\ \forall\, v \in S \\ y^\mathsf{T} b > 0 \end{array} \right)$$

holds.

Proof of Farkas

$$\left(\exists\, y \in \mathbb{R}^m,\, \beta \geq 0\ :\ \begin{array}{l} y^\intercal v \leq \beta,\ \forall\, v \in S \\ y^\intercal b > 0 \end{array} \right)$$

The value of $\beta \geq 0$ may be strictly positive, but we argue that it can be tightened to 0. Note that $S = \{Ax \,|\, x \geq 0\}$ has the property that $v \in S$ and $\alpha > 0$ implies $\alpha v \in S$. If $y^\intercal v > 0$ for any $v \in S$, then $y^\intercal(\alpha v) \to \infty$ and this would contradict the condition that $y^\intercal(\alpha v) \leq \beta$ for $(\alpha v) \in S$. Therefore, $y^\intercal v \leq 0$ for any $v \in S$, and we conclude that there exists $y \in \mathbb{R}^m$ such that

$$\left(\exists\, y \in \mathbb{R}^m\ :\ \begin{array}{l} y^\intercal v \leq 0,\ \forall\, v \in S \\ y^\intercal b > 0 \end{array} \right)$$

Proof of Farkas

$$\left(\exists\, y \in \mathbb{R}^m \;:\; \begin{array}{c} y^\intercal v \le 0,\ \forall\, v \in S \\ y^\intercal b > 0 \end{array}\right)$$

Next, plugging $S = \{Ax \,|\, x \ge 0\}$ into the condition above, we get

$$\left(\exists\, y \in \mathbb{R}^m \;:\; \begin{array}{c} y^\intercal Ax \le 0,\ \forall\, x \ge 0 \\ y^\intercal b > 0 \end{array}\right)$$

As discussed in a previous lemma,

$$\sup_{x \in \mathbb{R}^n_+} y^\intercal Ax = \begin{cases} 0 & \text{if } A^\intercal y \le 0 \\ \infty & \text{otherwise.} \end{cases}$$

(So $[y^\intercal Ax \le 0$ for all $x \ge 0]$ if and only if $A^\intercal y \le 0$.) Therefore,

$$\left(\exists\, y \in \mathbb{R}^m \;:\; \begin{array}{c} A^\intercal y \le 0 \\ y^\intercal b > 0 \end{array}\right)$$

Thus we conclude the second statement, and we conclude the proof. $\square$

Theorem (Strong duality).

Consider the primal and dual LPs

$$\begin{array}{lll} \underset{x\in\mathbb{R}^n}{\text{minimize}} & c^\intercal x & \\ \text{subject to} & Ax = b & \text{(P)} \\ & x \geq 0 & \end{array} \quad \overset{\text{dual}}{\longleftrightarrow} \quad \begin{array}{lll} \underset{y\in\mathbb{R}^m}{\text{maximize}} & b^\intercal y & \\ \text{subject to} & A^\intercal y \leq c, & \text{(D)} \end{array}$$

where $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, and $c \in \mathbb{R}^n$. Then, there are 4 (and no other) possible scenarios:

1. *(P) and (D) both infeasible ($-\infty = d_\star < p_\star = \infty$)*
2. *(P) unbounded and (D) infeasible ($-\infty = d_\star = p_\star$)*
3. *(P) infeasible and (D) unbounded ($d_\star = p_\star = \infty$)*
4. *(P) and (D) have solutions and s.d. holds ($-\infty < d_\star = p_\star < \infty$).*

In case 1, strong duality fails. In Cases 2–4, strong duality holds.

Proof of strong duality

Proof. Regarding feasibility, there are 4 cases:

	Primal Feasible	Dual Feasible
Case 1	✗	✗
Case 2	✓	✗
Case 3	✗	✓
Case 4	✓	✓

Case 1. There is nothing to show in this case.

Case 2. Primal feasible and dual infeasible. So, $-\infty = d_\star \leq p_\star < \infty$. It remains to show that $p_\star = -\infty$, i.e., we need to show the existence of a primal direction of unboundedness. The argument is similar to that of Case 3, and we leave it as a homework exercise.

Proof of strong duality

Case 3. Primal infeasible and dual feasible. So, $-\infty < d_\star \le p_\star = \infty$. It remains to show that $d_\star = \infty$, i.e., we need to show the existence of a dual direction of unboundedness.

$$\begin{array}{ll} \underset{y\in\mathbb{R}^m}{\text{maximize}} & b^\mathsf{T} y \\ \text{subject to} & A^\mathsf{T} y \le c \end{array} \tag{D}$$

Let $y_0 \in \mathbb{R}^m$ be a dual feasible point. Since the primal problem is infeasible, i.e., there is no x such that $[Ax = b \text{ and } x \ge 0]$, Farkas' lemma tells us that there is a y such that $[A^\mathsf{T} y \le 0 \text{ and } b^\mathsf{T} y > 0]$. Then,

$$\begin{aligned} & A^\mathsf{T}(y_0 + \alpha y) \le A^\mathsf{T} y_0 \le c && ((y_0+\alpha y) \text{ is feasible for } \alpha \ge 0) \\ & b^\mathsf{T}(y_0 + \alpha y) = b^\mathsf{T} y_0 + \alpha b^\mathsf{T} y \to \infty && (\text{objective is unbounded}) \end{aligned}$$

as $\alpha \to \infty$. (I.e., with a feasible point and a direction of unboundedness, we can drive the objective function to ∞.) Therefore, $d_\star = \infty$.

Proof of strong duality

Consider case 4. Primal and dual are feasible. So, $-\infty < d_\star \leq p_\star < \infty$. It remains to show that $p_\star = d_\star$.

Since the primal LP is feasible, i.e., there is an x such that [$Ax = b$ and $x \geq 0$]. By Farkas' lemma, we know that there is no y such that [$A^\intercal y \leq 0$ and $b^\intercal y > 0$].

Proof of strong duality

Let $v \in \mathbb{R}$. Then, by Farkas' lemma

$$\underbrace{\left(\exists\, x \in \mathbb{R}^n \;:\; \begin{aligned} Ax &= b \\ c^\intercal x &\le v \\ x &\ge 0 \end{aligned}\right)}_{=\text{there is a primal feasible } x \text{ with objective value no worse than } v} \quad \Leftrightarrow \quad \left(\exists\, x \in \mathbb{R}^n,\, s \in \mathbb{R} \;:\; \begin{array}{c} \begin{bmatrix} A & 0 \\ c^\intercal & 1 \end{bmatrix} \begin{bmatrix} x \\ s \end{bmatrix} = \begin{bmatrix} b \\ v \end{bmatrix} \\ x \ge 0,\; s \ge 0 \end{array}\right)$$

XOR

$$\left(\exists\, \tilde{y} \in \mathbb{R}^m,\, \tilde{\eta} \in \mathbb{R} \;:\; \begin{array}{c} \begin{bmatrix} A^\intercal & c \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \tilde{y} \\ \tilde{\eta} \end{bmatrix} \le \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ b^\intercal \tilde{y} + v\tilde{\eta} > 0 \end{array}\right)$$

Proof of strong duality

$$\left(\exists\, \tilde{y} \in \mathbb{R}^m,\, \tilde{\eta} \in \mathbb{R} \;:\; \begin{matrix} \begin{bmatrix} A^\intercal & c \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \tilde{y} \\ \tilde{\eta} \end{bmatrix} \le \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ b^\intercal \tilde{y} + v\tilde{\eta} > 0 \end{matrix}\right)$$

$$\overset{(1)}{\Leftrightarrow} \quad \left(\exists\, \tilde{y} \in \mathbb{R}^m,\, \eta \in \mathbb{R} \;:\; \begin{matrix} A^\intercal \tilde{y} \le \eta c \\ \eta \ge 0 \\ b^\intercal \tilde{y} > v\eta \end{matrix}\right)$$

$$\overset{(2)}{\Leftrightarrow} \quad \left(\exists\, \tilde{y} \in \mathbb{R}^m,\, \eta > 0 \;:\; \begin{matrix} A^\intercal \tilde{y} \le \eta c \\ b^\intercal \tilde{y} > v\eta \end{matrix}\right)$$

$$\overset{(3)}{\Leftrightarrow} \quad \left(\exists\, y \in \mathbb{R}^m \;:\; \begin{matrix} A^\intercal y \le c \\ b^\intercal y > v \end{matrix}\right) = \text{there is a dual feasible } y \text{ with objective value strictly better than } v$$

where (1) follows from setting $\tilde{\eta} = -\eta$, (2) follows from recognizing that $\eta \neq 0$ because we established in the previous slide that there is no y such that $A^\intercal y \le 0$ and $b^\intercal y > 0$, and (3) follows setting $y = \tilde{y}/\eta$.

Proof of strong duality

Therefore,

$$\underbrace{\left(\exists\, x \in \mathbb{R}^n \,:\, \begin{array}{r} Ax = b \\ c^\intercal x \le v \\ x \ge 0 \end{array}\right)}_{=\substack{\text{there is a primal feasible } x \text{ with} \\ \text{objective value no worse than } v}} \quad \text{XOR} \quad \underbrace{\left(\exists\, y \in \mathbb{R}^m \,:\, \begin{array}{r} A^\intercal y \le c \\ b^\intercal y > v \end{array}\right)}_{=\substack{\text{there is a dual feasible } y \text{ with} \\ \text{objective value strictly better than } v}}$$

Set $v = p_\star - \varepsilon$ with any $\varepsilon > 0$, note that such an x does not exist because a primal feasible x cannot attain an objective value better than $p_\star$. Since the XOR characterization, such a y does exist. So there is a dual feasible y attaining objective value $b^\intercal y > p_\star - \varepsilon$, and $p_\star - \varepsilon < d_\star \le p_\star$. By taking $\varepsilon \to 0$, we conclude $d_\star = p_\star$, i.e., strong duality holds.

Proof of strong duality

It remains to show that a primal and dual solution exists, i.e., we must show that the optimal value is attained.

By setting $v = d_\star = p_\star$, we see that

$$\left(\exists\, y \in \mathbb{R}^m \;:\; \begin{array}{r} A^\intercal y \leq c \\ b^\intercal y > d_\star \end{array}\right)$$

is fails, so

$$\left(\exists\, x \in \mathbb{R}^n \;:\; \begin{array}{r} Ax = b \\ c^\intercal x \leq p_\star \\ x \geq 0 \end{array}\right)$$

is holds. In particular, there is a x that is primal feasible and $c^\intercal x = p_\star$, so a primal solution exists.

The argument that a dual solution exists follows similar steps, and we leave it as a homework exercise. $\square$

Optimality conditions

The optimality conditions for constrained optimization are also called the Karush–Kuhn–Tucker (KKT) conditions.

Theorem (KKT conditions).

Let $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$, and $x \in \mathbb{R}^n$. The following are equivalent:

- *x solves*

$$\begin{array}{ll} \underset{x \in \mathbb{R}^n}{\text{minimize}} & c^\mathsf{T} x \\ \text{subject to} & Ax = b \\ & x \geq 0. \end{array}$$

- *There exists a $y \in \mathbb{R}^m$ such that x and y satisfy*

$$Ax = b, \quad x \geq 0, \quad A^\mathsf{T} y \leq c, \quad c^\mathsf{T} x = b^\mathsf{T} y.$$

- *There exists a $y \in \mathbb{R}^m$ such that x and y satisfy*

$$Ax = b, \quad x \geq 0, \quad A^\mathsf{T} y \leq c, \quad (c - A^\mathsf{T} y)^\mathsf{T} x = 0.$$

Optimality conditions

Theorem (KKT conditions, abridged).

The following are equivalent:

- x *solves the standard form LP.*
- *There exists a* $y \in \mathbb{R}^m$ *such that* x *and* y *satisfy*

$$Ax = b, \quad x \geq 0, \quad A^\mathsf{T} y \leq c, \quad c^\mathsf{T} x = b^\mathsf{T} y.$$

- *There exists a* $y \in \mathbb{R}^m$ *such that* x *and* y *satisfy*

$$Ax = b, \quad x \geq 0, \quad A^\mathsf{T} y \leq c, \quad (c - A^\mathsf{T} y)^\mathsf{T} x = 0.$$

The individual conditions are referred to as follows:

- $Ax = b$, $x \geq 0$ is called primal feasibility.
- $A^\mathsf{T} y \leq c$ is called dual feasibility.
- $c^\mathsf{T} x = b^\mathsf{T} y$ is called zero duality gap ($d_\star = p_\star$).
- $(c - A^\mathsf{T} y)^\mathsf{T} x$ is called complementary slackness.

Complementary slackness

The other conditions are somewhat self-explanatory.
What does complementary slackness mean?

If there exists a $y \in \mathbb{R}^m$ such that

$$Ax = b, \quad x \geq 0, \quad A^\intercal y \leq c, \quad \underbrace{(c - A^\intercal y)^\intercal}_{\geq 0} \underbrace{x}_{\geq 0} = 0,$$

then complementary slackness implies

$$(c - A^\intercal y)_i = 0 \quad \text{OR} \quad x_i = 0$$

for all $i = 1, \ldots, n$. (It is possible for both to be zero.)

So, the slackness (non-zeroness) across the coordinates do not overlap (is complementary).

In fact, the method of multipliers (a necessary but not sufficient condition for non-convex optimization) yields the same conditions. The theorem is saying that the conditions are necessary and sufficient for LPs.

Proof of optimality conditions

Theorem (KKT conditions, abridged).

The following are equivalent:

(a) x *solves the standard form LP.*

(b) *There exists a* $y \in \mathbb{R}^m$ *such that* x *and* y *satisfy*

$$Ax = b, \quad x \geq 0, \quad A^\mathsf{T} y \leq c, \quad c^\mathsf{T} x = b^\mathsf{T} y.$$

Proof. ((a)⇒(b)) If x is a solution, then $-\infty < p_\star < \infty$. This is Case 4 of the strong duality theorem, and this means there is a dual solution y and strong duality holds. So,

$$b^\mathsf{T} y = d_\star = p_\star = c^\mathsf{T} x$$

and we conclude (b).

((b)⇒(a)) By weak duality,

$$b^\mathsf{T} y \leq d_\star \leq p_\star \leq c^\mathsf{T} x = b^\mathsf{T} y,$$

and equality holds throughout by assumption of (b). In particular, $p_\star = c^\mathsf{T} x$ and we conclude that x is optimal for the primal LP.

Proof of optimality conditions

Theorem (KKT conditions, abridged).

The following are equivalent:

(b) *There exists a $y \in \mathbb{R}^m$ such that x and y satisfy*

$$Ax = b, \quad x \geq 0, \quad A^\intercal y \leq c, \quad c^\intercal x = b^\intercal y.$$

(c) *There exists a $y \in \mathbb{R}^m$ such that x and y satisfy*

$$Ax = b, \quad x \geq 0, \quad A^\intercal y \leq c, \quad (c - A^\intercal y)^\intercal x = 0.$$

((b)⇔(c)) Assume $Ax = b$, $x \geq 0$, and $A^\intercal y \leq c$. Then,

$$c^\intercal x = b^\intercal y \quad \Leftrightarrow \quad c^\intercal x = y^\intercal A x \quad \Leftrightarrow \quad (c - A^\intercal y)^\intercal x = 0.$$

□

Outline

LP applications

LP theory

Weak duality

Strong duality

Examples

Simplex method in tableau form

Visualize separating hyperplane

Let $\beta \in \mathbb{R}$ and $y = (1, 2) \in \mathbb{R}^2$. Visualize the separating hyperplane

$$H = \{x \mid y^\intercal x \leq \beta\}.$$

Show visualizations

Visualizing a primal LP

Consider the

$$\begin{array}{ll} \underset{x \in \mathbb{R}^2}{\text{minimize}} & c^\intercal x \\ \text{subject to} & [1, -1]\, x \leq 1 \\ & [1, -1]\, x \geq -2 \\ & x \geq 0. \end{array}$$

Show visualization of the feasible set.

Visualizing a primal LP

Consider the

$$\begin{array}{ll} \underset{x \in \mathbb{R}^2}{\text{minimize}} & c^\intercal x \\ \text{subject to} & [1, -1]\, x \leq 1 \\ & [1, -1]\, x \geq -2 \\ & x \geq 0. \end{array}$$

with

$$c = \begin{bmatrix} \cos(\theta) \\ \sin(\theta) \end{bmatrix}.$$

The LP is bounded for $\theta \in (-\frac{\pi}{4}, \frac{3\pi}{4})$.

Show visualization of solution.

Dual solution as an algebraic proof of optimality

Let $c = (2, -1)$. Then we can see that $x_\star = (0, 2)$ solves the problem

$$\begin{array}{ll} \underset{x \in \mathbb{R}^2}{\text{minimize}} & c^\intercal x \\ \text{subject to} & [1, -1]\, x \leq 1 \\ & [-1, 1]\, x \leq 2 \\ & [-1, 0]\, x \leq 0 \\ & [0, -1]\, x \leq 0 \end{array}$$

with $p_\star = -2$.

The dual LP has the form

$$\begin{array}{ll} \underset{y \in \mathbb{R}^4}{\text{maximize}} & -y_1 - 2y_2 \\ \text{subject to} & -y_1 + y_2 + y_3 = 2 \\ & y_1 - y_2 + y_4 = -1 \\ & y \geq 0 \end{array}$$

and has solution $y_\star = (0, 1, 1, 0)$.

Dual solution as an algebraic proof of optimality

In the problem

$$\begin{array}{ll} \underset{x\in\mathbb{R}^2}{\text{minimize}} & [2,-1]\,x \\ \text{subject to} & [1,-1]\,x \le 1 \\ & [-1,1]\,x \le 2 \\ & [-1,0]\,x \le 0 \\ & [0,-1]\,x \le 0, \end{array}$$

which has solution $x_\star = (0,2)$, the "active" constraints are $[-1,1]\,x \le 2$ and $[-1,0]\,x \le 0$. We can visually see that those are the constraints that prevent us from further reducing the objective value.

Then, combining inequalities

$$1 \cdot ([1,-1]\,x \ge -2) + 1 \cdot \big([1,0]\,x \ge 0\big) = ([2,-1]\,x \ge -2),$$

and therefore $p_\star \ge -2$.

Physics interpretation of the dual solution

$$\begin{array}{ll} \underset{x\in\mathbb{R}^2,\, t\in\mathbb{R}}{\text{minimize}} & \underbrace{[2,-1]}_{=c^\intercal}\, x \\ \text{subject to} & [1,-1]\, x \le 1 \\ & [-1,1]\, x \le 2 \\ & [-1,0]\, x \le 0 \\ & [0,-1]\, x \le 0, \end{array}$$

Imagine a ball positioned at x and we "push" it with force $-c$. The ball is constrained to be within the walls defined by the inequality constraints.

The ball will stop moving once we reach $x_\star = (0, 2)$.

Once the ball is stationary, the walls must be exerting forces that add up to c, exactly countering the force $-c$ that we are exerting. Only the walls touching the ball can exert force.

The force a wall exerts is normal (perpendicular) to the surface and must point inward. If wall $([-1,1]\, x \le 2)$ exerts force $(1,-1)$ and wall $([-1,0]\, x \le 0)$ exerts force $(1,0)$, this adds up to c.

Case with $-\infty = d_\star < p_\star = \infty$

In the strong duality theorem, we stated that there are 4 cases. Precisely, we showed that nothing outside of the 4 cases are possible. But can all 4 cases really occur?

The previous example provides instances of Case 2 and 4.

Case 3 is possible, similar to Case 2.

Is Case 1 possible? Yes, it is. In the standard for LP, consider

$$A = [0], \quad b = 1, \quad c = -1.$$

Then, the primal and dual standard form LPs are both infeasible:

$$\begin{array}{ll} \underset{x \in \mathbb{R}}{\text{minimize}} & -x \\ \text{subject to} & 0 \cdot x = 1 \\ & x \geq 0 \end{array} \quad \overset{\text{dual}}{\longleftrightarrow} \quad \begin{array}{ll} \underset{y \in \mathbb{R}}{\text{maximize}} & y \\ \text{subject to} & 0 \cdot y \leq -1 \end{array}$$

Outline

Simplex method

Key observation: Solutions of LPs always occur on "corner" points, which are called *basic feasible points*.

The *simplex method* traverses basic feasible points to find a solution.

We will describe the simplex method through an example demonstration.

Gaussian elimination operations

Lemma.

Let

$$A = \begin{bmatrix} a_1^\intercal \\ \vdots \\ a_m^\intercal \end{bmatrix} \in \mathbb{R}^{m \times n}, \qquad b = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} \in \mathbb{R}^m.$$

For any $\alpha \neq 0$ *and* $i \in \{1, \dots, m\}$,

$$\left\{ x \in \mathbb{R}^n \ : \ \begin{bmatrix} a_1^\intercal \\ \vdots \\ a_i^\intercal \\ \vdots \\ a_m^\intercal \end{bmatrix} x = \begin{bmatrix} b_1 \\ \vdots \\ b_i \\ \vdots \\ b_m \end{bmatrix} \right\} = \left\{ x \in \mathbb{R}^n \ : \ \begin{bmatrix} a_1^\intercal \\ \vdots \\ \alpha a_i^\intercal \\ \vdots \\ a_m^\intercal \end{bmatrix} x = \begin{bmatrix} b_1 \\ \vdots \\ \alpha b_i \\ \vdots \\ b_m \end{bmatrix} \right\}$$

I.e., scaling a row of a linear system by a nonzero scalar does not change the set of x's satisfying the linear system.

Gaussian elimination operations

Lemma.

$$\underbrace{\left\{ x \in \mathbb{R}^n \;:\; \begin{bmatrix} a_1^\mathsf{T} \\ \vdots \\ a_i^\mathsf{T} \\ \vdots \\ a_m^\mathsf{T} \end{bmatrix} x = \begin{bmatrix} b_1 \\ \vdots \\ b_i \\ \vdots \\ b_m \end{bmatrix} \right\}}_{=S_1} = \underbrace{\left\{ x \in \mathbb{R}^n \;:\; \begin{bmatrix} a_1^\mathsf{T} \\ \vdots \\ \alpha a_i^\mathsf{T} \\ \vdots \\ a_m^\mathsf{T} \end{bmatrix} x = \begin{bmatrix} b_1 \\ \vdots \\ \alpha b_i \\ \vdots \\ b_m \end{bmatrix} \right\}}_{=S_2}$$

Proof. If $x \in S_1$, then $a_j^\mathsf{T} x = b_j$ for $j = 1, \ldots, m$. By scaling the i-th row by α, we show that x satisfies the linear system of S_2, i.e., $x \in S_2$. If $x \in S_2$, we follow the same argument and scale the i-th row by $1/\alpha$ to conclude that $x \in S_1$. Thus, $S_1 = S_2$. $\square$

Gaussian elimination operations

Lemma.

Let

$$A = \begin{bmatrix} a_1^\intercal \\ \vdots \\ a_m^\intercal \end{bmatrix} \in \mathbb{R}^{m \times n}, \qquad b = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} \in \mathbb{R}^m.$$

For any $i, j \in \{1, \ldots, m\}$ *such that* $i \neq j$,

$$\left\{ x \in \mathbb{R}^n \ : \ \begin{bmatrix} a_1^\intercal \\ \vdots \\ a_i^\intercal \\ \vdots \\ a_m^\intercal \end{bmatrix} x = \begin{bmatrix} b_1 \\ \vdots \\ b_i \\ \vdots \\ b_m \end{bmatrix} \right\} = \left\{ x \in \mathbb{R}^n \ : \ \begin{bmatrix} a_1^\intercal \\ \vdots \\ (a_i - a_j)^\intercal \\ \vdots \\ a_m^\intercal \end{bmatrix} x = \begin{bmatrix} b_1 \\ \vdots \\ b_i - b_j \\ \vdots \\ b_m \end{bmatrix} \right\}$$

I.e., subtracting the j-th row from the i-th row does not change the set of x's satisfying the linear system.

Gaussian elimination operations

Lemma.

$$\underbrace{\left\{ x \in \mathbb{R}^n \; : \; \begin{bmatrix} a_1^\intercal \\ \vdots \\ a_i^\intercal \\ \vdots \\ a_m^\intercal \end{bmatrix} x = \begin{bmatrix} b_1 \\ \vdots \\ b_i \\ \vdots \\ b_m \end{bmatrix} \right\}}_{=S_1} = \underbrace{\left\{ x \in \mathbb{R}^n \; : \; \begin{bmatrix} a_1^\intercal \\ \vdots \\ (a_i - a_j)^\intercal \\ \vdots \\ a_m^\intercal \end{bmatrix} x = \begin{bmatrix} b_1 \\ \vdots \\ b_i - b_j \\ \vdots \\ b_m \end{bmatrix} \right\}}_{=S_2}$$

Proof. If $x \in S_1$, then we can subtract $a_j^\intercal x = b_j$ from $a_i^\intercal x = b_i$ to get $(a_i - a_j)^\intercal x = (b_i - b_j)$. So, $x \in S_2$. On the other hand, if $x \in S_2$, then we can add $a_j^\intercal x = b_j$ to $(a_i - a_j)^\intercal x = (b_i - b_j)$ to recover $a_i^\intercal x = b_i$. So, $x \in S_1$. Thus, $S_1 = S_2$. $\square$

Simplex method: Transformation to standard form

Consider the problem

$$\begin{array}{ll} \underset{x\in\mathbb{R}^3}{\text{minimize}} & -2x_1 - 3x_2 - 4x_3 \\ \text{subject to} & 3x_1 + 2x_2 + x_3 \le 10 \\ & 2x_1 + 5x_2 + 3x_3 \le 15 \\ & x \ge 0. \end{array}$$

This is equivalent to

$$\begin{array}{ll} \underset{z\in\mathbb{R},\, x\in\mathbb{R}^3_+,\, s\in\mathbb{R}^2_+}{\text{minimize}} & z \\ \text{subject to} & -2x_1 - 3x_2 - 4x_3 = z \\ & 3x_1 + 2x_2 + x_3 + s_1 = 10 \\ & 2x_1 + 5x_2 + 3x_3 + s_2 = 15. \end{array}$$

Note, z has no positivity constraint. We express $x \ge 0$, $s \ge 0$ implicitly for the sake of brevity.

Simplex method: Basic feasible point

This is equivalent to

$$\begin{array}{ll} \underset{z\in\mathbb{R},\, x\in\mathbb{R}^3_+,\, s\in\mathbb{R}^2_+}{\text{minimize}} & z \\ \text{subject to} & \left[\begin{array}{cccccc} 1 & 2 & 3 & 4 & 0 & 0 \\ \hline 0 & 3 & 2 & 1 & 1 & 0 \\ 0 & 2 & 5 & 3 & 0 & 1 \end{array}\right] \begin{bmatrix} z \\ x_1 \\ x_2 \\ x_3 \\ s_1 \\ s_2 \end{bmatrix} = \left[\begin{array}{c} 0 \\ \hline 10 \\ 15 \end{array}\right] \end{array}$$

There are $p = 2$ equality constraints, excluding the one defining z. This system has a *basic feasible point*

$$(z, x_1, x_2, x_3, s_1, s_2) = (0, 0, 0, 0, 10, 15).$$

To have a basic feasible point, we need the linear system to have p columns corresponding to $e_2, e_3, \ldots, e_{p+1}$ (almost forming an identity matrix block), and the basic feasible point has only p nonzeros, excluding z, corresponding to those p columns.

The p nonzero variables s_1 and s_2 are called *basic variables*.

Simplex method: Pivot operation

To improve the objective, we will change the set of basic variables. This operation is called a *pivot operation*.

We can increase x_1, x_2, x_3 and that would reduce the objective value, because 2, 3, and 4 are positive.

$$\begin{aligned} \underset{z\in\mathbb{R},\, x\in\mathbb{R}^3_+,\, s\in\mathbb{R}^2_+}{\text{minimize}} \quad & z \\ \text{subject to} \quad & \left[\begin{array}{cccccc} 1 & 2 & 3 & 4 & 0 & 0 \\ \hline 0 & 3 & 2 & 1 & 1 & 0 \\ 0 & 2 & 5 & 3 & 0 & 1 \end{array}\right] \begin{bmatrix} z \\ x_1 \\ x_2 \\ x_3 \\ s_1 \\ s_2 \end{bmatrix} = \left[\begin{array}{c} 0 \\ \hline 10 \\ 15 \end{array}\right] \end{aligned}$$

Let's pick x_3, since the coefficient 4 is the largest. (z will be reduced more rapidly since the coefficient is large.)

This choice makes x_3 the *entering variable*.

Simplex method: Pivot operation

The set of basic variables must be maintained at size p. Since x_3 is chosen to be an entering variable, we must choose a *leaving variable* among s_1 and s_2. Which one shall we choose?

$$\begin{array}{ll} \underset{z\in\mathbb{R},\, x\in\mathbb{R}^3_+,\, s\in\mathbb{R}^2_+}{\text{minimize}} & z \\ \text{subject to} & \left[\begin{array}{cccccc} 1 & 2 & 3 & \color{blue}{4} & 0 & 0 \\ \hline 0 & 3 & 2 & 1 & \color{red}{1} & 0 \\ 0 & 2 & 5 & 3 & 0 & \color{red}{1} \end{array}\right] \begin{bmatrix} z \\ x_1 \\ x_2 \\ \color{blue}{x_3} \\ \color{red}{s_1} \\ \color{red}{s_2} \end{bmatrix} = \left[\begin{array}{c} 0 \\ \hline 10 \\ 15 \end{array}\right] \end{array}$$

As we update x_3, z will move accordingly. We keep the other non-basic variables (0-variables) fixed. We need to choose one of the basic variable to be the leaving variable (make it 0). Which one?

$$(z, x_1, x_2, x_3, s_1, s_2) = (0, 0, 0, 0, \color{red}{10}, \color{red}{15}) \mapsto \left\{ \begin{array}{c} (*, 0, 0, \color{red}{*}, 0, \color{red}{*}) \\ \text{or} \\ (*, 0, 0, \color{red}{*}, \color{red}{*}, 0) \end{array} \right.$$

Simplex method: Minimum ratio test

We can choose the leaving variable via the *minimum ratio test*.
For educational purposes, however, let us try both.

If we choose s_1 as the leaving variable, we will do the Gaussian elimination operation to make the column of x_3 be e_2, replacing the column of s_1:

$$\begin{aligned} \underset{z\in\mathbb{R},\, x\in\mathbb{R}^3_+,\, s\in\mathbb{R}^2_+}{\text{minimize}} \quad & z \\ \text{subject to} \quad & \left[\begin{array}{cccccc} 1 & 2 & 3 & \color{blue}{4} & 0 & 0 \\ \hline 0 & 3 & 2 & 1 & \color{red}{1} & 0 \\ 0 & 2 & 5 & 3 & 0 & \color{red}{1} \end{array}\right] \begin{bmatrix} z \\ x_1 \\ x_2 \\ \color{blue}{x_3} \\ \color{red}{s_1} \\ \color{red}{s_2} \end{bmatrix} = \left[\begin{array}{c} 0 \\ \hline 10 \\ 15 \end{array}\right] \end{aligned}$$

This is equivalent to

$$\begin{aligned} \underset{z\in\mathbb{R},\, x\in\mathbb{R}^3_+,\, s\in\mathbb{R}^2_+}{\text{minimize}} \quad & z \\ \text{subject to} \quad & \left[\begin{array}{cccccc} 1 & -10 & -5 & 0 & -4 & 0 \\ \hline 0 & 3 & 2 & \color{red}{1} & 1 & 0 \\ 0 & -7 & -1 & 0 & -3 & \color{red}{1} \end{array}\right] \begin{bmatrix} z \\ x_1 \\ x_2 \\ \color{red}{x_3} \\ s_1 \\ \color{red}{s_2} \end{bmatrix} = \left[\begin{array}{c} -40 \\ \hline 10 \\ -15 \end{array}\right] \end{aligned}$$

Simplex method: Minimum ratio test

$$\begin{array}{ll} \underset{z\in\mathbb{R},\, x\in\mathbb{R}^3_+,\, s\in\mathbb{R}^2_+}{\text{minimize}} & z \\ \text{subject to} & \left[\begin{array}{cccccc} 1 & -10 & -5 & 0 & -4 & 0 \\ \hline 0 & 3 & 2 & 1 & 1 & 0 \\ 0 & -7 & -1 & 0 & -3 & 1 \end{array}\right] \begin{bmatrix} z \\ x_1 \\ x_2 \\ x_3 \\ s_1 \\ s_2 \end{bmatrix} = \left[\begin{array}{c} -40 \\ \hline 10 \\ -15 \end{array}\right] \end{array}$$

Then, our corresponding basic feasible point is

$$(z, x_1, x_2, x_3, s_1, s_2) = (-40, 0, 0, 10, 0, -15)$$

The objective value improved $0 \mapsto -40$, but the basic variable $s_2 = -15$ violates the constraint $s_2 \geq 0$. This wasn't the right choice.

Simplex method: Minimum ratio test

Instead, choose s_2 as the leaving variable:

$$\begin{array}{ll} \underset{z\in\mathbb{R},\, x\in\mathbb{R}^3_+,\, s\in\mathbb{R}^2_+}{\text{minimize}} & z \\ \text{subject to} & \left[\begin{array}{cccccc} 1 & 2 & 3 & 4 & 0 & 0 \\ \hline 0 & 3 & 2 & 1 & 1 & 0 \\ 0 & 2 & 5 & 3 & 0 & 1 \end{array}\right] \begin{bmatrix} z \\ x_1 \\ x_2 \\ x_3 \\ s_1 \\ s_2 \end{bmatrix} = \left[\begin{array}{c} 0 \\ \hline 10 \\ 15 \end{array}\right] \end{array}$$

Scaling the last row, we get

$$\begin{array}{ll} \underset{z\in\mathbb{R},\, x\in\mathbb{R}^3_+,\, s\in\mathbb{R}^2_+}{\text{minimize}} & z \\ \text{subject to} & \left[\begin{array}{cccccc} 1 & 2 & 3 & 4 & 0 & 0 \\ \hline 0 & 3 & 2 & 1 & 1 & 0 \\ 0 & 2/3 & 5/3 & 1 & 0 & 1/3 \end{array}\right] \begin{bmatrix} z \\ x_1 \\ x_2 \\ x_3 \\ s_1 \\ s_2 \end{bmatrix} = \left[\begin{array}{c} 0 \\ \hline 10 \\ 5 \end{array}\right] \end{array}$$

Simplex method: Minimum ratio test

Performing Gaussian elimination operation to make the column of x_3 be e_3, replacing the column of s_2, we get

$$\begin{array}{ll} \underset{z\in\mathbb{R},\, x\in\mathbb{R}^3_+,\, s\in\mathbb{R}^2_+}{\text{minimize}} & z \\ \text{subject to} & \left[\begin{array}{cccccc} 1 & -2/3 & -11/3 & 0 & 0 & -4/3 \\ \hline 0 & 7/3 & 1/3 & 0 & 1 & -1/3 \\ 0 & 2/3 & 5/3 & 1 & 0 & 1/3 \end{array}\right] \begin{bmatrix} z \\ x_1 \\ x_2 \\ x_3 \\ s_1 \\ s_2 \end{bmatrix} = \left[\begin{array}{c} -20 \\ \hline 5 \\ 5 \end{array}\right] \end{array}$$

Then, our corresponding basic feasible point is

$$(z, x_1, x_2, x_3, s_1, s_2) = (-20, 0, 0, 5, 5, 0)$$

The objective value improved $0 \mapsto -20$, and the inequality constraints are respected.

Simplex method: Minimum ratio test

The *minimum ratio test* allows one to choose whether s_1 or s_2 should be the leaving variables without trial and error.

$$\begin{array}{ll} \underset{z\in\mathbb{R},\, x\in\mathbb{R}^3_+,\, s\in\mathbb{R}^2_+}{\text{minimize}} & z \\ \text{subject to} & \left[\begin{array}{cccccc} 1 & 2 & 3 & 4 & 0 & 0 \\ \hline 0 & 3 & 2 & 1 & 1(\text{✗}) & 0 \\ 0 & 2 & 5 & 3 & 0 & 1(\checkmark) \end{array}\right] \begin{bmatrix} z \\ x_1 \\ x_2 \\ x_3 \\ s_1 \\ s_2 \end{bmatrix} = \left[\begin{array}{c} 0 \\ \hline 10 \\ 15 \end{array}\right] \qquad \begin{array}{l} \\ (10/1 = 10\ \text{✗}) \\ (15/3 = 5\ \checkmark) \end{array} \end{array}$$

We compute the ratio of the RHS over the column corresponding to the entering variable x_3. The leaving variable is the basic variable corresponding to the row with the minimum ratio value.

Proof of optimality

$$\begin{array}{ll} \underset{z\in\mathbb{R},\, x\in\mathbb{R}^3_+,\, s\in\mathbb{R}^2_+}{\text{minimize}} & z \\ \text{subject to} & \left[\begin{array}{cccccc} 1 & -2/3 & -11/3 & 0 & 0 & -4/3 \\ \hline 0 & 7/3 & 1/3 & 0 & \color{red}{1} & -1/3 \\ 0 & 2/3 & 5/3 & \color{red}{1} & 0 & 1/3 \end{array}\right] \begin{bmatrix} z \\ x_1 \\ x_2 \\ \color{red}{x_3} \\ \color{red}{s_1} \\ s_2 \end{bmatrix} = \left[\begin{array}{c} -20 \\ \hline 5 \\ 5 \end{array}\right] \end{array}$$

$$(z, x_1, x_2, x_3, s_1, s_2) = (-20, 0, 0, \color{red}{5}, \color{red}{5}, 0)$$

Recall that the set of feasible points of an LP is convex.

At this point, all of the coefficients of the non-basic (zero) variables are negative, and this tells us that our basic feasible point is globally optimal. If we change $(x_1, x_2, x_3, s_1, s_2)$ by $(\delta_1, \delta_2, \delta_3, \eta_1, \eta_2)$, then

$$(z, x_1, x_2, x_3, s_1, s_2) = (-20+\tfrac{2}{3}\delta_1+2\delta_2+\tfrac{4}{3}\eta_2, 0+\delta_1, 0+\delta_2, 5+\delta_3, 5+\eta_1, 0+\eta_2),$$

where $\delta_1, \delta_2, \eta_2 \geq 0$ to respect the nonnegativity constraints. Therefore, we cannot reduce z while remaining feasible, and we conclude optimality.