Math 373:
Introduction to Optimization

John C. Bowman and Pouria Ramazi
University of Alberta
Edmonton, Canada

December 6, 2025

Contents

Preface

The figures in this text were drawn with the vector graphics language `Asymptote` (freely available at `http://asymptote.sourceforge.net`).

Chapter 1

Linear Programming

[Bertsimas & Tsitsiklis 1997]

Optimization is an important area of applied mathematics that is widely used in many areas of society, including industry, business, science, and economics. Some optimization problems involve an *infinite number* of variables (like those that arise in the calculus of variations that you may encounter in a mechanics course). However, in this course, we focus on optimization problems that involve maximizing or minimizing a function of a *finite number* of variables subject to *finitely* many inequality (or equality) constraints. In particular, we will examine the case where both the function and the constraints are linear functions of their arguments. This subject is often confusingly called *linear programming*, even though it it deals more with linear algebra and geometry than the technical aspects of computer programming. The term "programming" is used here in the sense of detailed military logistics planning, stemming from the work of George Danzig in the US Air Force and at the RAND Corporation just after the end of World War II. The alternative term *linear optimization* is perhaps a more suitable name.

To illustrate what linear programming is all about, let us begin with a simple *optimization problem*:

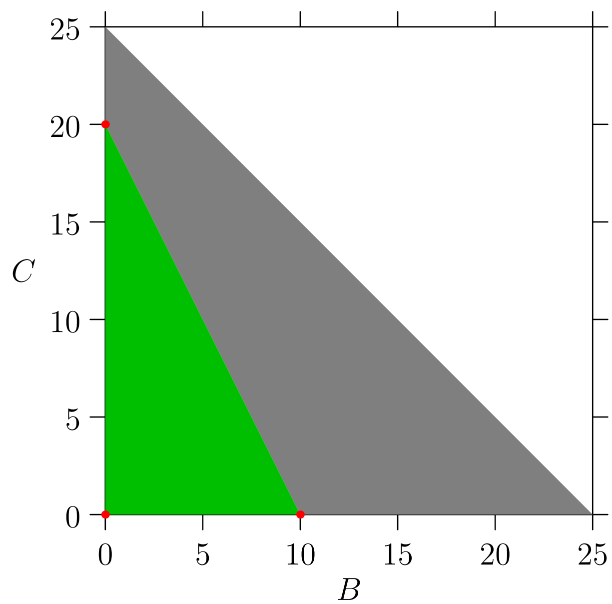
- A farmer has 25 hectares of land on which to grow barley or corn. He earns \$600 for each hectare of barley and \$500 for each hectare of corn. Harvesting the barley requires 20 hours of labour per hectare, and harvesting the corn requires 10 hours of labour per hectare. The farmer can afford 200 hours of labour. How can he maximize his profit?

To help us answer this question, let us adopt units of hectares, dollars, and hours. We first define the *decision variables* that express the choices facing the farmer: how many hectares of barley to grow (B) and how many hectares of corn (C) to grow. We are told that the profit is $f(B,C) = 600B + 500C$.

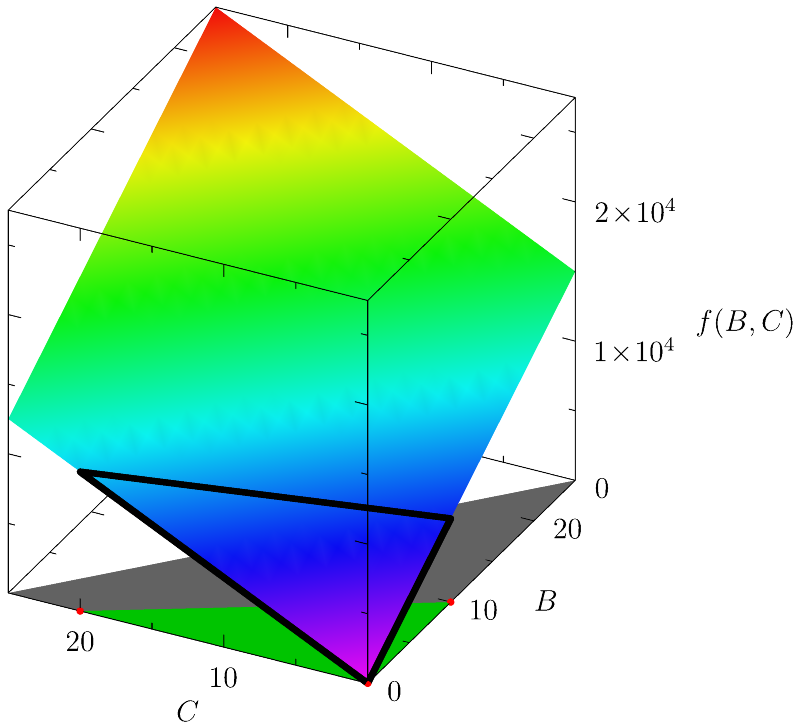
There are four *constraints* in the problem. First, the total number of hectares is $B + C \leq 25$. Second, the total amount of labour is limited: $20B + 10C \leq 200$.

Finally, since you cannot plant negative hectares of grain, we know that $B \geq 0$ and $C \geq 0$.

Remark: It is helpful to illustrate this linear programming problem graphically. In the following graph, the shaded (green+gray) region in the B–C plane depicts the constraints $B + C \leq 25$, $B \geq 0$, $C \geq 0$, which is further refined by the constraint $20B + 10C \leq 200$ to the green triangle. The green triangle represents the *feasible set*, which is the collection of all *feasible solutions* (B, C) that satisfy the four given constraints.



Remark: Let us now consider the behaviour of the *objective function* $f(B, C) = 600B + 500C$ as B and C vary over the box $[0, 25] \times [0, 25]$ in the B–C plane.



One notices that the function takes on its maximum value over the green region at the vertex $(0, 20)$ and its minimum value at the vertex $(0, 0)$. The graph illustrates that the solution to the farmer's maximization problem occurs at $B = 0$ and $C = 20$. That is, in order to maximize his profit, the farmer should only plant 20 hectares of corn; this will use up all 200 hours of his labour but reward him with a tidy profit of $10 000.

Remark: Notice that the extrema of the objective function occur at vertices of the feasible set.

Remark: Another feature that we notice in this example is the first constraint $B + C \leq 25$ was superfluous and can be removed from the problem. In some applications, the removal of such superfluous constraints can make solving linear programming problems much more efficient!

Remark: An optimization problem to an *objective function* f of n *decision variables* $\boldsymbol{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$, over the *feasible set* $F \subset \mathbb{R}^n$ of solutions satisfying all of the constraints, can be written symbolically as

$$\max_{\boldsymbol{x} \in F} f(\boldsymbol{x})$$

or

$$\min_{\boldsymbol{x} \in F} f(\boldsymbol{x}).$$

- For example, in the problem

$$\max_{\substack{x_1, x_2 \\ x_1^2 + x_2^2 \leq 1}} \; (x_1^2 + x_2^2),$$

  the feasible set is the unit disk. This is an example of a *nonlinear optimization problem*: both the objective function $f(x_1, x_2) = x_1^2 + x_2^2$ and constraint $x_1^2 + x_2^2 \leq 1$ are nonlinear in the decision variables x_1 and x_2.

Definition: The *optimal value* is the desired extreme (maximum or minimum) value (if it exists) of the objective function.

Definition: A feasible solution $\boldsymbol{x}^*$ that extremizes the objective function is called an *optimal feasible solution*, or simply an *optimal solution*.

Definition: The *optimal set* is the set of optimal solutions; that is, the set of feasible solutions at which the objective function f takes on its optimal value (if it exists).

Remark: In the above problem, the optimal value is 1 and the set of optimal solutions is the unit circle.

Remark: In determining optimal solutions, one considers only points $\boldsymbol{x}$ that are feasible.

- For the problem

$$\max_{\substack{x_1^2 + x_2^2 \leq 1 \\ x_2 \geq 0}} (x_1^2 + x_2^2),$$

  the optimal set is the upper half circle.

- For the problem

$$\max_{\substack{x \geq -1 \\ x \leq 1}} (x^2 + 2),$$

  the feasible set is $[-1, 1]$, the optimal value is 3, and $x = -1$ and $x = 1$ are the optimal solutions.

- For the problem

$$\min_{\substack{x \geq -1 \\ x \leq 3 \\ y \leq 2}} (x^2 + y^2),$$

  the feasible set is $\{(x, y) : x \in [-1, 3], y \in (-\infty, 2]\}$, the optimal value is 0 and the optimal solution is $(0, 0)$.

- For the problem

$$\max_{x>0} \frac{1}{x},$$

the feasible set is $(0, \infty)$; we say that the optimal value is ∞, but there are no optimal solutions (the optimal set is empty).

- For the problem

$$\min_{x>0} \frac{1}{x},$$

the feasible set is again $(0, \infty)$; we say that the optimal value is 0, but there are still no optimal solutions.

Remark: There is no need to study maximization and minimization problems separately: maximizing f is equivalent to minimizing $-f$.

Remark: In typical real-life problems, determining the optimal set is often even more important that knowing the optimal value. For example, if you are asked to solve an optimization problem to maximize profit, it doesn't help much to know the optimal value if you don't know how to achieve it!

Remark: Solving an optimization problem means finding the optimal value and **all** optimal solutions.

Remark: Various outcomes to an optimization problem are possible:

$$\text{The problem may be} \begin{cases} \text{infeasible (no feasible solutions)} \\ \text{feasible} \begin{cases} \text{no optimal solutions} \\ \text{optimal solutions} \begin{cases} \text{finitely many} \\ \text{infinitely many} \end{cases} \end{cases} \end{cases}$$

Problem 1.1: Suppose you are asked to produce two kinds of meals: economy and deluxe. The economy meal sells for \$3/kg and the deluxe version sells for \$4/kg. The ingredients are rice and lamb. The economy version should contain at most 25% lamb (by weight), while the deluxe version should contain at least 50% lamb. The cost of rice is \$1/kg, whereas lamb costs \$2/kg. You have 300 kg of rice and 100 kg of lamb available. Formulate a linear programming problem that determines how much rice and lamb should you put into each dish to maximize your profit, assuming that all of the prepared dishes sell.

Let the decision variables r_e and r_d represent the weight of rice in the economy and deluxe dishes respectively, and ℓ_e and ℓ_d represent the corresponding weights of lamb. The objective function is $f(r_e, r_d, \ell_e, \ell_d) = 3(r_e + \ell_e) + 4(r_d + \ell_d) - (r_e + r_d) - 2(\ell_e + \ell_d)$. One wants to maximize this function subject to the constraints

$$r_e + r_d \leq 300,$$

$$\ell_e + \ell_d \leq 100,$$

$$\ell_e \leq \frac{1}{4}(r_e + \ell_e),$$

$$\ell_d \geq \frac{1}{2}(r_d + \ell_d),$$

and

$$r_e, \ell_e, r_d, \ell_d \geq 0.$$

1.A. Linear programming problems

Definition: A function $f : \mathbb{R}^n \to \mathbb{R}$ is *linear* if it can be expressed as

$$f(x_1, \ldots, x_n) = \sum_{i=1}^{n} c_i x_i,$$

for some real numbers c_i.

Remark: Equivalently, a function $f : \mathbb{R}^n \to \mathbb{R}$ is *linear* if

$$f(\alpha \boldsymbol{x} + \beta \boldsymbol{y}) = \alpha f(\boldsymbol{x}) + \beta f(\boldsymbol{y})$$

for all vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ and real numbers α and β.

Definition: A *linear constraint* on the variables $x_1, \ldots, x_n$ has one of the following forms, where $a_1, \ldots, a_n$ and b are real numbers:

- linear equalities, such as
$$\sum_{i=1}^{n} a_i x_i = b;$$
- linear inequalities, such as
$$\sum_{i=1}^{n} a_i x_i \geq b.$$

Definition: Although for $n > 1$, $\mathbb{R}^n$ is not an ordered field, it is nevertheless possible to introduce a *partial ordering* on $\mathbb{R}^n$ component-wise: given two vectors $\boldsymbol{x} = (x_1, \dots, x_n)$ and $\boldsymbol{y} = (y_1, \dots, y_n)$ where $x_i \le y_i$ for all i, we say $\boldsymbol{x} \le \boldsymbol{y}$. Similarly, $\boldsymbol{x} \ge \boldsymbol{y}$ means that $x_i \ge y_i$ for all i. Corresponding definitions hold for strict inequalities, with $\le$ replaced by $<$ and $\ge$ replaced by $>$.

Definition: In a *linear programming problem*, one seeks to minimize a *linear cost function* $\boldsymbol{c}^\mathsf{T}\boldsymbol{x} = \sum_{j=1}^n c_j x_j$ over all vectors $\boldsymbol{x} = (x_1, \dots, x_n)$ in $\mathbb{R}^n$, subject to a set of linear equality and inequality constraints. The vector $\boldsymbol{c} = (c_1, \dots, c_n)$ is called a *cost vector* and we denote its transpose by $\boldsymbol{c}^\mathsf{T}$. Note that $\boldsymbol{c}^\mathsf{T}\boldsymbol{x} = \boldsymbol{c}\cdot\boldsymbol{x}$.

That is, a linear programming problem is of the form

$$\begin{aligned} \text{minimize} \quad & \boldsymbol{c}^\mathsf{T}\boldsymbol{x} \\ \text{subject to} \quad & \boldsymbol{a}_i^\mathsf{T}\boldsymbol{x} \le b_i, \quad i \in \mathcal{M}_1, \\ & \boldsymbol{a}_i^\mathsf{T}\boldsymbol{x} \ge b_i, \quad i \in \mathcal{M}_2, \\ & \boldsymbol{a}_i^\mathsf{T}\boldsymbol{x} = b_i, \quad i \in \mathcal{M}_3, \\ & \boldsymbol{x}_j \le 0, \quad j \in \mathcal{N}_1, \\ & \boldsymbol{x}_j \ge 0, \quad j \in \mathcal{N}_2, \end{aligned} \tag{1.1}$$

where $\mathcal{M}_1$, $\mathcal{M}_2$, $\mathcal{M}_3$, $\mathcal{N}_1$, and $\mathcal{N}_2$ are finite sets of integers, and the variables $\boldsymbol{x} = (x_1, \dots, x_n)$ are called *decision variables*. When $\boldsymbol{x}$ satisfies all of the given constraints, it is called a *feasible solution* or *feasible vector*. If there are no restrictions on the sign of x_j, i.e., j does not belong to $\mathcal{N}_1 \cup \mathcal{N}_2$, we say that x_j is a *free* or *unrestricted* variable. The linear function $f(\boldsymbol{x}) = \boldsymbol{c}^\mathsf{T}\boldsymbol{x}$ is called the *objective function* or *cost function*. A feasible solution $\boldsymbol{x}^*$ that minimizes the objective function is called an *optimal feasible solution* or *optimal solution*. The value $\boldsymbol{c}^\mathsf{T}\boldsymbol{x}^*$ is the *optimal cost*.

Problem 1.2: By introducing an appropriate *coefficient matrix* $\boldsymbol{A}$ and *target vector* $\boldsymbol{b}$, show that (1.1) can be rewritten in the compact form

$$\begin{aligned} \text{minimize} \quad & \boldsymbol{c}^\mathsf{T}\boldsymbol{x} \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} \ge \boldsymbol{b}. \end{aligned}$$

1.B. Standard form

A linear programming problem of the form

$$\begin{aligned} \text{minimize} \quad & \boldsymbol{c}^\mathsf{T}\boldsymbol{x} \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \\ & \boldsymbol{x} \ge \boldsymbol{0}. \end{aligned}$$

is said to be in *standard form*.

Remark: If we let $\boldsymbol{A}_1, \dots, \boldsymbol{A}_n$ represent the column vectors of the matrix $\boldsymbol{A}$, the constraint $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ can be written as

$$\sum_{j=1}^{n} \boldsymbol{A}_j x_j = \boldsymbol{b}.$$

The goal is to synthesize the target vector $\boldsymbol{b}$ from a nonnegative amount x_j of each resource $\boldsymbol{A}_j$, such that the cost $\sum_{j=1}^{n} c_j x_j$ is minimized, where c_j is the unit cost of the jth resource.

- **The diet problem.** Suppose that there are n different foods and m different nutrients. The following table indicates the nutritional content of a unit of each food:

	food 1	$\cdots$	food n
nutrient 1	a_{11}	$\cdots$	a_{1n}
$\vdots$	$\vdots$		$\vdots$
nutrient m	a_{m1}	$\cdots$	a_{mn}

Let $\boldsymbol{A}$ be the $m \times n$ matrix with entries a_{ij}, so that its jth column $\boldsymbol{A}_j$ represents the nutritional content of the jth food. Let $\boldsymbol{b}$ be the vector expressing the requirements of an ideal diet. The standard-form problem aims to mix nonnegative quantities x_j of the available foods to achieve an ideal diet at minimal cost. The condition for synthesizing the ideal diet is

$$\sum_{j=1}^{n} \boldsymbol{A}_j x_j = \boldsymbol{b}.$$

If $\boldsymbol{b}$ instead represents the minimal requirements of an adequate diet, the constraint becomes

$$\sum_{j=1}^{n} \boldsymbol{A}_j x_j \geq \boldsymbol{b}.$$

1.C. Reduction to standard form

Using the following procedures, it is possible to reduce any linear programming problem to standard form:

1. *Elimination of free variables*: Each free variable x_j can be replaced by the difference $x_j^+ - x_j^-$ of two additional nonnegative variables x_j^+ and x_j^- (every real number can be expressed as the difference of two nonnegative numbers). This removes x_j from the problem but adds two new constraints: $x_j^+ \geq 0$ and $x_j^- \geq 0$.

2. *Elimination of inequality constraints*: Given an inequality constraint of the form

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

we introduce a nonnegative *slack variable* s_i to obtain the standard-form constraints

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

$$s_i \geq 0.$$

Likewise, for each constraint of the form

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

we introduce a nonnegative *surplus variable* s_i, such that

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

$$s_i \geq 0.$$

Since they work similarly, we will often refer to slack and surplus variables collectively as (generic) slack variables.

- For the linear programming problem

$$\begin{array}{lrcrl} \text{minimize} & x_1 & + & 2x_2 & \\ \text{subject to} & x_1 & + & 3x_2 & \leq 2, \\ & 2x_1 & + & x_2 & = 1, \\ & & & x_1 & \geq 0, \end{array}$$

an equivalent problem in standard form is

$$\begin{array}{lrcrcrl} \text{minimize} & x_1 & + & 2x_2^+ - 2x_2^- & & & \\ \text{subject to} & x_1 & + & 3x_2^+ - 3x_2^- & + & x_3 & = 2, \\ & 2x_1 & + & x_2^+ - x_2^- & & & = 1, \\ & & & & & x_1 & \geq 0, \\ & & & & & x_2^+ & \geq 0, \\ & & & & & x_2^- & \geq 0, \\ & & & & & x_3 & \geq 0, \end{array}$$

For example, for the feasible solution $(x_1, x_2) = (1, -1)$ to the original problem, the standard-form problem has a corresponding feasible solution $(x_1, x_2^+, x_2^-, x_3) = (1, 0, 1, 4)$ (not unique), with the same cost. Conversely, the feasible solution $(x_1, x_2^+, x_2^-, x_3) = (2, 1, 4, 9)$ to the standard-form problem corresponds to the solution $(x_1, x_2) = (2, -3)$ to the original problem at the same cost.

1.D. Strict inequalities

Linear programming problems are typically formulated with nonstrict inequalities. Nevertheless, a problem involving strict inequality constraints such as

$$\boldsymbol{a}^\intercal \boldsymbol{x} > b$$

can be handled by first solving the linear programming problem obtained by replacing the above inequality with the nonstrict version

$$\boldsymbol{a}^\intercal \boldsymbol{x} \geq b$$

and then checking whether an optimal solution $\boldsymbol{x}^*$ to this modified problem satisfies $\boldsymbol{a}^\intercal \boldsymbol{x}^* > b$. If it does not, then an optimal feasible solution does not exist, but there exist feasible solutions arbitrarily close to $\boldsymbol{x}^*$.

Chapter 2

The Geometry of Linear Programming

2.A. Polyhedra

Definition: A *convex polyhedron* is a set that can be described in the form

$$\{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b}\},$$

where $\boldsymbol{A}$ is an $m \times n$ matrix and $\boldsymbol{b}$ is a vector in $\mathbb{R}^m$.

Remark: By the arguments used to reduce a linear programming problem to standard form, a convex polyhedron can be equivalently defined as a set

$$\{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq \boldsymbol{0}\}.$$

Remark: A convex polyhedron can either extend to infinity or can be contained in a finite region.

Definition: A set $\mathcal{S} \subset \mathbb{R}^n$ is *bounded* if there exists some constant K such that $|\boldsymbol{x}| \leq K$ for every element $\boldsymbol{x} \in \mathcal{S}$. Typically, we will use the Euclidean norm $|\boldsymbol{x}| = \sqrt{\sum_{i=1}^n x_i^2}$.

Definition: Let $\boldsymbol{a} \in \mathbb{R}^n, \boldsymbol{a} \neq \boldsymbol{0}$ and $b \in \mathbb{R}$. Then

1. The set $\{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{a}^\intercal \boldsymbol{x} = b\}$ is called a *hyperplane*.
2. The set $\{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{a}^\intercal \boldsymbol{x} \geq b\}$ is called a *half space*.

Notice that the hyperplane is the boundary of the corresponding half space and is perpendicular to $\boldsymbol{a}$.

Remark: In $\mathbb{R}^3$, the xy plane ($z=0$) is the hyperplane $\boldsymbol{a}\cdot\boldsymbol{x}=0$ corresponding to $\boldsymbol{a}=(0,0,1)$ and $b=0$. Everything lying on or above that ($z\geq 0$) is a half space. Everything lying on or below that ($z\leq 0$) is another half space. Thus, the hyperplane $z=0$ divides the space $\mathbb{R}^3$ into two half spaces whose union makes up all of $\mathbb{R}^3$.

Remark: A convex polyhedron can thus be expressed as an intersection of half spaces.

Definition: The *graph* of a function $f:\mathbb{R}^n\to\mathbb{R}$ is the set of ordered pairs

$$\{(\boldsymbol{x},f(\boldsymbol{x})):\boldsymbol{x}\in\mathbb{R}^n\}.$$

Remark: The graph of f is just the set of values taken on by the function

$$F:\mathbb{R}^n\to\mathbb{R}^{n+1},\quad F(\boldsymbol{x})=(\boldsymbol{x},f(\boldsymbol{x})).$$

2.B. Review of convexity for functions $f:\mathbb{R}\to\mathbb{R}$

Definition: A function $f:\mathbb{R}\to\mathbb{R}$ is *convex* (sometimes called *concave up*) on an interval $I\in\mathbb{R}$ if its graph on every subinterval $[p,q]$ of I lies below or on the secant line segment joining $(p,f(p))$ and $(q,f(q))$.

Definition: A function f is *concave* (sometimes called *concave down*) on an interval I if $-f$ is convex on I.

Remark: Since the equation of the line through $(p,f(p))$ and $(q,f(q))$ is

$$y=f(p)+\frac{f(q)-f(p)}{q-p}(x-p),$$

the definition of convex says

$$f(x)\leq f(p)+\frac{f(q)-f(p)}{q-p}(x-p)\quad\text{for all } x\in[p,q],\quad p,q\in I. \tag{2.1}$$

This condition may be rewritten by re-expressing the *linear interpolation* of f between p and q on the right-hand side of (2.1):

$$f(x)\leq\left(\frac{q-x}{q-p}\right)f(p)+\left(\frac{x-p}{q-p}\right)f(q)\quad\text{for all } x\in[p,q],\quad p,q\in I. \tag{2.2}$$

It is sometimes convenient to introduce the *parameter* $t=\dfrac{q-x}{q-p}$, in terms of which we may express $x=q-(q-p)t$ and

$$\frac{x-p}{q-p}=\frac{(q-p)-(q-p)t}{q-p}=1-t.$$

This allow us to restate (2.2) in *parametric form*:

$$f(tp+(1-t)q)\leq tf(p)+(1-t)f(q)\quad\text{for all } t\in[0,1],\quad p,q\in I. \tag{2.3}$$

2.C. Convexity for functions $f : \mathbb{R}^n \to \mathbb{R}$

Definition: Given points $\boldsymbol{p}, \boldsymbol{q} \in \mathbb{R}^n$, the set $\{t\boldsymbol{p} + (1-t)\boldsymbol{q} : t \in [0,1]\}$ describes the *line segment* with endpoints $\boldsymbol{p}$ and $\boldsymbol{q}$.

Definition: A set $S \subset \mathbb{R}^n$ is *convex* if for every $\boldsymbol{p}, \boldsymbol{q} \in S$ the line segment joining $\boldsymbol{p}$ and $\boldsymbol{q}$ is contained entirely within S.

Definition: A function $f : \mathbb{R}^n \to \mathbb{R}$ is *convex* if for every $\boldsymbol{p}, \boldsymbol{q}$ and $t \in [0,1]$ we have $f(t\boldsymbol{p} + (1-t)\boldsymbol{q}) \leq tf(\boldsymbol{p}) + (1-t)f(\boldsymbol{q})$.

Remark: Equivalently, a function is *convex* if its graph lies below or on the secant line segment joining $(\boldsymbol{p}, f(\boldsymbol{p}))$ and $(\boldsymbol{q}, f(\boldsymbol{q}))$.

Definition: A function $f : \mathbb{R}^n \to \mathbb{R}$ is *concave* if for every $\boldsymbol{p}, \boldsymbol{q}$ and $t \in [0,1]$ we have $f(t\boldsymbol{p} + (1-t)\boldsymbol{q}) \geq tf(\boldsymbol{p}) + (1-t)f(\boldsymbol{q})$.

Definition: An *affine* function $f : \mathbb{R}^n \to \mathbb{R}$ has the form $f(\boldsymbol{x}) = \boldsymbol{a}^\intercal \boldsymbol{x} + c$ for some constant c.

Remark: The only functions that are both convex and concave are affine functions.

Remark: A real-valued function f has a *local minimum* at $\boldsymbol{p}$ if $f(\boldsymbol{q}) \geq f(\boldsymbol{p})$ for all $\boldsymbol{q}$ sufficiently near $\boldsymbol{p}$.

Remark: A real-valued function f has a *global minimum* at $\boldsymbol{p}$ if $f(\boldsymbol{q}) \geq f(\boldsymbol{p})$ for all $\boldsymbol{q}$.

Remark: For convex functions on convex sets, the following theorem establishes that local minima are also global minima.

Theorem 2.1: *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function and let $S \subset \mathbb{R}^n$ be a convex set. A local minimum of f at a point of S is actually a global minimum over S.*

Proof: Given a local minimum m of f at $\boldsymbol{x} \in S$, suppose that there exists a point $\boldsymbol{y} \in S$ where $f(\boldsymbol{y}) < m$. Then since f is convex,

$$f(t\boldsymbol{x} + (1-t)\boldsymbol{y}) \leq tf(\boldsymbol{x}) + (1-t)f(\boldsymbol{y}) < tm + (1-t)m = m$$

for every $t \in [0,1)$, noting that $t\boldsymbol{x} + (1-t)\boldsymbol{y}$ lies in the convex set S. As $t \to 1^-$ we then see that f has values less than m at points arbitrarily close to $\boldsymbol{x}$, contradicting the fact that f has a local minimum at $\boldsymbol{x}$. Thus there are no such points $\boldsymbol{y}$. That is, the local minimum of f at $\boldsymbol{x}$ is actually a global minimum.

2.D. The convex hull

Definition: Let $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k \in \mathbb{R}^n$, $t_1, \ldots, t_k \geq 0$, and $\sum_{i=1}^k t_i = 1$. Then

1. The vector $\sum_{i=1}^k t_i \boldsymbol{x}_i$ is a **convex combination** of the vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$.
2. The *convex hull* of the vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$ is the set of all convex combinations of these vectors. Equivalently, it is the smallest convex polyhedron that contains the points $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$.

Remark: Every half space $H = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{a}^\intercal \boldsymbol{x} \geq b\}$ is convex: let $\boldsymbol{x}, \boldsymbol{y} \in H$. For every $t \in [0,1]$ we see that $\boldsymbol{a}^\intercal(t\boldsymbol{x} + (1-t)\boldsymbol{y}) \geq tb + (1-t)b = b$. That is, $t\boldsymbol{x} + (1-t)\boldsymbol{y} \in H$. Thus H is convex.

Theorem 2.2 (Properties of Convex Sets):

(i) *The intersection of convex sets is convex.*

(ii) *Every convex polyhedron* $\{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b} \in \mathbb{R}^m\}$ *is a convex set.*

(iii) *A convex combination of a finite number of elements of a convex set also belongs to that set.*

(iv) *The convex hull of a finite number of points is a convex set.*

Proof:

(i) Consider a collection of convex sets S_i, where i belongs to some index set I. If $\boldsymbol{x}, \boldsymbol{y} \in \bigcap_{i \in I} S_i$ then for each $i \in I$ and $t \in [0,1]$, we know from the convexity of S_i that $t\boldsymbol{x} + (1-t)\boldsymbol{y} \in S_i$. Therefore $t\boldsymbol{x} + (1-t)\boldsymbol{y} \in \bigcap_{i \in I} S_i$. Thus $\bigcap_{i \in I} S_i$ is convex.

(ii) Since a convex polyhedron is an intersection of half spaces, which we have seen are convex, this follows from (i).

(iii) By the definition of convexity, a convex combination of two elements of a convex set S lies in that set. Now suppose that every convex combination of $k \geq 2$ elements of S belongs to S. Given a nontrivial convex combination $\sum_{i=1}^{k+1} t_i \boldsymbol{x}_i$ of $k+1$ elements $\boldsymbol{x}_i$ of S, where $t_i \in [0,1)$ for $i = 1, \ldots, k+1$, such that $\sum_{i=1}^{k+1} t_i = 1$, we express

$$\sum_{i=1}^{k+1} t_i \boldsymbol{x}_i = t_{k+1}\boldsymbol{x}_{k+1} + (1 - t_{k+1}) \sum_{i=1}^{k} \frac{t_i}{1 - t_{k+1}} \boldsymbol{x}_i.$$

The summation in the second term belongs to S since it is a convex combination of k elements of S: $\sum_{i=1}^{k} t_i = \sum_{i=1}^{k+1} t_i - t_{k+1} = 1 - t_{k+1}$. Thus $\sum_{i=1}^{k+1} t_i \boldsymbol{x}_i$ is a convex combination of two elements of S and is therefore also in S. By induction, a convex combination of any finite number of elements of a convex set also belongs to that set.

(iv) Let $\boldsymbol{y} = \sum_{i=1}^{k} \alpha_i \boldsymbol{x}_i$ and $\boldsymbol{z} = \sum_{i=1}^{k} \beta_i \boldsymbol{x}_i$ be two elements of the convex hull S of points $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$, where the non-negative weights α_i and β_i satisfy $\sum_{i=1}^{k} \alpha_i = \sum_{i=1}^{k} \beta_i = 1$. Then for $t \in [0, 1]$,

$$t\boldsymbol{y} + (1-t)\boldsymbol{z} = t\sum_{i=1}^{k} \alpha_i \boldsymbol{x}_i + (1-t)\sum_{i=1}^{k} \beta_i \boldsymbol{x}_i = \sum_{i=1}^{k} [t\alpha_i + (1-t)\beta_i]\boldsymbol{x}_i$$

is a convex combination of $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$, noting that

$$\sum_{i=1}^{k} [t\alpha_i + (1-t)\beta_i] = t\sum_{i=1}^{k} \alpha_i + (1-t)\sum_{i=1}^{k} \beta_i = t + (1-t) = 1.$$

That is, $t\boldsymbol{y} + (1-t)\boldsymbol{z}$ belongs to S. We conclude that S is a convex set.

2.E. Piecewise linear convex objective functions

Definition: A function $f : \mathbb{R}^n \to \mathbb{R}$ is *piecewise linear* if it is linear on each of a finite number of intervals of $\mathbb{R}^n$.

- The absolute value function

$$|x| = \begin{cases} x & \text{if } x \geq 0, \\ -x & \text{if } x < 0 \end{cases}$$

is piecewise linear.

Remark: Although we stated earlier that we will focus on objective functions that are linear, one can easily generalize the methods we will develop to functions that are *piecewise linear*. We will then be able to apply linear programming to problems of the form

$$\begin{aligned} &\text{minimize} \quad \max_{i=1,\ldots,m} (\boldsymbol{c}_i^{\mathsf{T}}\boldsymbol{x} + d_i) \\ &\text{subject to} \quad \boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b}. \end{aligned}$$

Since $\max_{i=1,\ldots,m} (\boldsymbol{c}_i^{\mathsf{T}}\boldsymbol{x} + d_i)$ is the smallest number M such that $M \geq \boldsymbol{c}_i^{\mathsf{T}}\boldsymbol{x} + d_i$ for all i, the above optimization problem is equivalent to the linear programming problem

$$\begin{aligned} &\text{minimize} \quad M \\ &\text{subject to} \quad M \geq \boldsymbol{c}_i^{\mathsf{T}}\boldsymbol{x} + d_i, \qquad i = 1, \ldots, m, \\ &\qquad\qquad\quad \boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b}. \end{aligned}$$

Theorem 2.3 (Piecewise convex functions): *Let $f_1, \dots f_m : \mathbb{R}^n \to \mathbb{R}$ be convex functions. Then the function $f(\boldsymbol{x}) = \max_{i=1,\dots,m} f_i(\boldsymbol{x})$ is also convex.*

Proof: Let $t \in [0, 1]$. Since

$$\begin{aligned} f(t\boldsymbol{p} + (1-t)\boldsymbol{q}) &= \max_{i=1,\dots,m} f_i(t\boldsymbol{p} + (1-t)\boldsymbol{q}) \\ &\le \max_{i=1,\dots,m} [t f_i(\boldsymbol{p}) + (1-t) f_i(\boldsymbol{q})] \\ &\le t \max_{i=1,\dots,m} f_i(\boldsymbol{p}) + (1-t) \max_{i=1,\dots,m} f_i(\boldsymbol{q}) \\ &= t f(\boldsymbol{p}) + (1-t) f(\boldsymbol{q}). \end{aligned}$$

for all $\boldsymbol{p}$ and $\boldsymbol{q}$, we see that f is indeed convex.

- Since the absolute value function $f(x) = |x| = \max\{x, -x\}$, we see from Theorem 2.3 that the absolute value function is convex (as well as piecewise linear).

Remark: Piecewise linear convex functions can be used to approximate more general functions.

Remark: In addition to handling piecewise linear convex objective functions, we can also handle piecewise affine constraints like

$$\max_{i=1,\dots,m} (\boldsymbol{e}_i^\intercal \boldsymbol{x} + d_i) \le b$$

by rewriting them as separate constraints $\boldsymbol{e}_i^\intercal \boldsymbol{x} + d_i \le b$, $i = 1, \dots, m$.

Problem 2.1: Consider functions $f, g : [a, b] \to \mathbb{R}$. Is it true that

$$\max_{x \in [a,b]} [f(x) + g(x)] = \max_{x \in [a,b]} f(x) + \max_{x \in [a,b]} g(x)?$$

Prove or provide a counterexample.

The statement is false: consider for example $f(x) = x$ and $g(x) = 1 - x$ on $[0, 1]$. We see that $\max_{x \in [0,1]} [f(x) + g(x)] = 1$ but $\max_{x \in [0,1]} f(x) + \max_{x \in [0,1]} g(x) = 1 + 1 = 2$. We can only guarantee that

$$\max_{x \in [a,b]} [f(x) + g(x)] \le \max_{x \in [a,b]} f(x) + \max_{x \in [a,b]} g(x).$$

2.F. Problems involving absolute values

Consider problems of the form

$$\begin{aligned} \text{minimize} \quad & \sum_{j=1}^{n} c_j \, |x_j| \\ \text{subject to} \quad & \boldsymbol{Ax} \geq \boldsymbol{b}, \end{aligned}$$

where $\boldsymbol{x} = (x_1, \ldots, x_n)$ and the cost coefficients c_j are positive. Although the objection function here is easily shown to be piecewise linear and convex (being the sum of piecewise linear convex functions) and therefore can be handled in the manner just described, a more efficient method is available. We can use the fact that $|x_j|$ is the smallest number M_j that is an upper bound for both x_j and $-x_j$ to rewrite the linear programming problem as:

$$\begin{aligned} \text{minimize} \quad & \sum_{j=1}^{n} c_j M_j \\ \text{subject to} \quad & \boldsymbol{Ax} \geq \boldsymbol{b}, \\ & x_j \leq M_j, \qquad j = 1, \ldots, n, \\ & -x_j \leq M_j, \qquad j = 1, \ldots, n. \end{aligned}$$

An alternative method is to replace each occurrence of $|x_j|$ with $x_j^+ + x_j^-$ and then each of the remaining occurrences of x_j with $x_j^+ - x_j^-$, where x_j^+ and x_j^- are new nonnegative decision variables:

$$\begin{aligned} \text{minimize} \quad & \sum_{j=1}^{n} c_j \left(x_j^+ + x_j^- \right) \\ \text{subject to} \quad & \boldsymbol{Ax}^+ - \boldsymbol{Ax}^- \geq \boldsymbol{b}, \\ & \boldsymbol{x}^+, \boldsymbol{x}^- \geq \boldsymbol{0}, \end{aligned}$$

where $\boldsymbol{x}^+ = (x_1^+, \ldots, x_n^+)$ and $\boldsymbol{x}^- = (x_1^-, \ldots, x_n^-)$. The equivalence of these two formulations follows from the observation that at an optimal solution, for each j one of x_j^+ or x_j^- must be zero (otherwise we could decrease both variables, preserving feasibility and reducing the cost), from which it follows that $|x_j| = x_j^+ + x_j^-$.

Remark: A similar argument shows that one can apply this technique to any linear programming problem of the form

$$\begin{aligned} \text{minimize} \quad & \sum_{j=1}^{n} c_j \, |x_j| + \sum_{j=1}^{n} d_j x_j \\ \text{subject to} \quad & \sum_{j=1}^{n} \boldsymbol{A}_j \, |x_j| + \sum_{j=1}^{n} \boldsymbol{B}_j x_j \leq \boldsymbol{b}, \end{aligned}$$

where each c_j and all entries in each $\boldsymbol{A}_j$ are nonnegative.

Problem 2.2:

Reformulate the problem

$$\begin{array}{ll} \text{minimize} & x_1 + |x_2 - 1| \\ \text{subject to} & |x_1 + 1| + |x_2| \leq 2. \end{array}$$

Let $x_1 + 1 = x_1^+ - x_1^-$ and replace $|x_1 + 1|$ by $x_1^+ + x_1^-$. Let $x_2 = x_2^+ - x_2^-$, replacing $|x_2|$ by $x_2^+ + x_2^-$. Finally, let $x_2 - 1 = x_3^+ - x_3^-$, replacing $|x_2 - 1|$ with $x_3^+ + x_3^-$. We obtain the equivalent linear programming problem

$$\begin{array}{lrl} \text{minimize} & x_1^+ - x_1^- - 1 + x_3^+ + x_3^- & \\ \text{subject to} & x_1^+ + x_1^- + x_2^+ + x_2^- & \leq 2, \\ & x_2^+ - x_2^- - x_3^+ + x_3^- & = 1, \\ & x_1^+, x_1^-, x_2^+, x_2^-, x_3^+, x_3^- & \geq 0. \end{array}$$

Note that we can ignore the additive constant -1 in the objective function for the purposes of finding optimal solutions (but we would need to account for this constant in computing the optimal cost).

2.G. Extreme points

Definition: Let P be a convex polyhedron. A point $\boldsymbol{x} \in P$ is an extreme point if it cannot be expressed as a convex combination of two other points of P. That is, a point $\boldsymbol{x} \in P$ is an *extreme point* of P if we **cannot** find two distinct points $\boldsymbol{y}, \boldsymbol{z} \in P$ and a scalar $t \in (0,1)$ such that $\boldsymbol{x} = t\boldsymbol{y} + (1-t)\boldsymbol{z}$.

Definition: Let P be a convex polyhedron. A vector $\boldsymbol{x} \in P$ is a vertex of P if P lies entirely on one side of some hyperplane $\{\boldsymbol{y} : \boldsymbol{c}^\intercal \boldsymbol{y} = \boldsymbol{c}^\intercal \boldsymbol{x}\}$, intersecting the hyperplane only at $\boldsymbol{x}$. That is $\boldsymbol{x} \in P$ is a *vertex* of P if there exists some $\boldsymbol{c}$ such that $\boldsymbol{c}^\intercal \boldsymbol{x} < \boldsymbol{c}^\intercal \boldsymbol{y}$ for all $\boldsymbol{y} \in P, \boldsymbol{y} \neq \boldsymbol{x}$.

Definition: Consider a polyhedron $P \in \mathbb{R}^n$ defined in terms of

$$\begin{array}{ll} \boldsymbol{a}_i^\intercal \boldsymbol{x} \geq b_i & i \in M_1, \\ \boldsymbol{a}_i^\intercal \boldsymbol{x} \leq b_i & i \in M_2, \\ \boldsymbol{a}_i^\intercal \boldsymbol{x} = b_i & i \in M_3. \end{array}$$

If a vector $\boldsymbol{x}$ satisfies $\boldsymbol{a}_i^\intercal \boldsymbol{x} = b_i$ for some i in M_1, M_2, or M_3, we say that the corresponding constraint is *active* at $\boldsymbol{x}$.

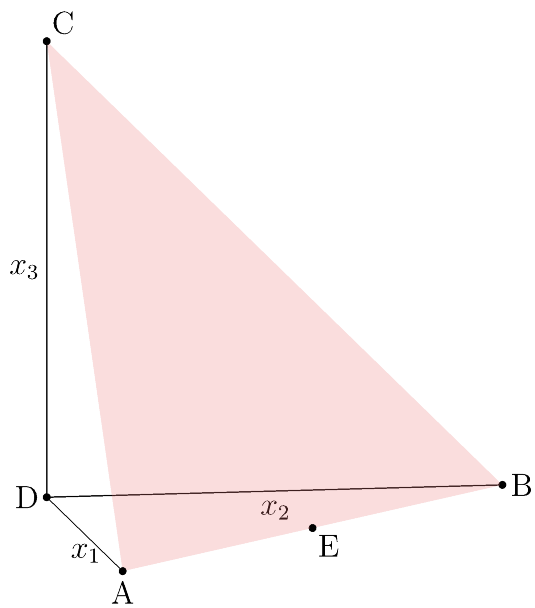


Figure 2.1: The polyhedron $\{(x_1, x_2, x_3) : x_1 + x_2 + x_3 = 1, x_1, x_2, x_3 \geq 0\}$.

- Given the polyhedron $\{(x_1, x_2, x_3) : x_1 + x_2 + x_3 = 1, x_1, x_2, x_3 \geq 0\}$ shown in Figure 2.1, which constraints are active at the points A, B, C, D, E?
  A: $x_1 + x_2 + x_3 = 1$, $x_2 = 0$, $x_3 = 0$
  B: $x_1 + x_2 + x_3 = 1$, $x_1 = 0$, $x_3 = 0$
  C: $x_1 + x_2 + x_3 = 1$, $x_1 = 0$, $x_2 = 0$
  D: $x_1 = 0$, $x_2 = 0$, $x_3 = 0$
  E: $x_1 + x_2 + x_3 = 1$, $x_3 = 0$

Theorem 2.4 (Active constraints): *Let $\boldsymbol{x} \in \mathbb{R}^n$ and $I = \{i : \mathbf{a}_i^\intercal \boldsymbol{x} = b_i\}$ be the set of indices of active constraints at $\boldsymbol{x}$. Then the following are equivalent:*

(i) *There exists n vectors in the set $\{\boldsymbol{a}_i : i \in I\}$ that are linearly independent.*

(ii) *The span of the vectors $\{\boldsymbol{a}_i : i \in I\}$ is $\mathbb{R}^n$.*

(iii) *The system of equations $\mathbf{a}_i^\intercal \boldsymbol{y} = b_i$, $i \in I$, has the unique solution $\boldsymbol{y} = \boldsymbol{x}$.*

(i) $\iff$ (ii): Suppose that n of the vectors $\boldsymbol{a}_i, i \in I$, are linearly independent. The subspace spanned by these vectors is n dimensional and is therefore $\mathbb{R}^n$. Conversely, suppose that the vectors $\{\boldsymbol{a}_i : i \in I\}$ span $\mathbb{R}^n$. Then n of these vectors form a basis of R^n and must therefore be linearly independent.

$\overline{\text{(iii)}} \iff \overline{\text{(ii)}}$: Suppose that $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are distinct solutions to the system of equations $\boldsymbol{a}_i^\intercal \boldsymbol{x} = b_i$, $i \in I$. Then $\boldsymbol{d} = \boldsymbol{x}_1 - \boldsymbol{x}_2 \in \mathbb{R}^n$ satisfies $\boldsymbol{a}_i^\intercal \boldsymbol{d} = 0$ for all $i \in I$. That is, $\boldsymbol{d} \neq \boldsymbol{0}$ is orthogonal to every $\boldsymbol{a}_i$, $i \in I$, and therefore cannot be expressed as a linear combination of these vectors. That is, $\{\boldsymbol{a}_i : i \in I\}$ do not span all of $\mathbb{R}^n$. Conversely, if the vectors $\boldsymbol{a}_i$, $i \in I$, do not span all of $\mathbb{R}^n$, choose a nonzero vector $\boldsymbol{d}$ orthogonal to the subspace they span, so that $\boldsymbol{a}_i^\intercal \boldsymbol{d} = 0$ for all $i \in I$. For every solution $\boldsymbol{x}$ to the system of equations $\boldsymbol{a}_i^\intercal \boldsymbol{x} = b_i$, $i \in I$, the vector $\boldsymbol{x} + \boldsymbol{d}$ will then be a distinct solution to the same system of equations.

Definition: Consider a convex polyhedron P defined by linear equality and inequality constraints, and let $\boldsymbol{x} \in \mathbb{R}^n$. Then the vector $\boldsymbol{x}$ is a *basic solution* if

1. all equality constraints are active;
2. there are n linearly independent active constraints at $\boldsymbol{x}$.

Definition: If $\boldsymbol{x}$ is a basic solution that satisfies all of the constraints, we say that it is a *basic feasible solution*.

Remark: A feasible solution $\boldsymbol{x}$ is basic iff n linearly independent constraints are active at $\boldsymbol{x}$.

- Given the polyhedron P shown in Figure 2.1, which of the points A, B, C, D, E are feasible, basic, or basic feasible solutions?
  A: basic feasible
  B: basic feasible
  C: basic feasible
  D: nonbasic infeasible
  E: nonbasic feasible

Q. What would happen if the equality constraint $x_1 + x_2 + x_3 = 1$ were to be replaced by the constraints $x_1 + x_2 + x_3 \leq 1$ and $x_1 + x_2 + x_3 \geq 1$?

A. The point D would become a basic solution (but still infeasible).

Remark: Whether a point is a basic solution or not may depend on the way that the polyhedron is represented!

- Given the polygon in Figure 2.2, which of the points A, B, C, D, E, F, G, H are basic, nonbasic, feasible, or basic feasible solutions?
  A: basic
  B: basic
  C: basic feasible
  D: basic feasible
  E: basic feasible
  F: basic feasible
  G: nonbasic infeasible
  H: nonbasic feasible

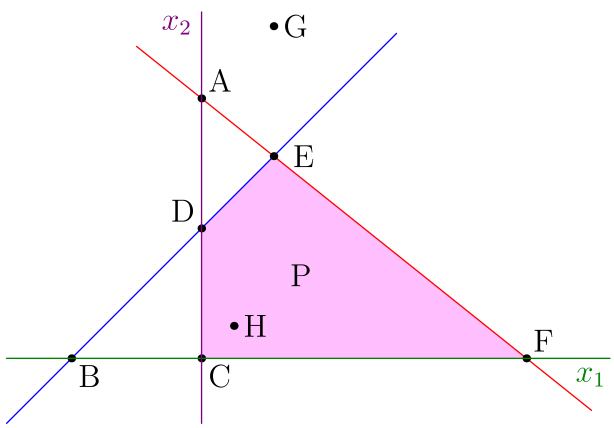


Figure 2.2: Examples of basic solutions.

Remark: So far, we have introduced two geometric definitions (extreme point, vertex) and one algebraic condition (basic feasible solution). The following theorem shows that these three definitions are actually equivalent, so that these concepts can be used interchangeably!

Theorem 2.5 (Characterization of Vertices): *Let P be a nonempty convex polyhedron and let $\boldsymbol{x} \in P$. Then the following are equivalent:*

(i) $\boldsymbol{x}$ *is a vertex;*

(ii) $\boldsymbol{x}$ *is an extreme point;*

(iii) $\boldsymbol{x}$ *is a basic feasible solution.*

Proof: Represent P as a set of equality constraints $\boldsymbol{a}_i^\intercal \boldsymbol{x} = b_i$ and inequality constraints $\boldsymbol{a}_i^\intercal \boldsymbol{x} \geq b_i$.

Vertex $\Rightarrow$ Extreme point:
Let $\boldsymbol{x} \in P$ be a vertex. Then there exists $\boldsymbol{c} \in \mathbb{R}^n$ such that $\boldsymbol{c}^\intercal \boldsymbol{x} < \boldsymbol{c}^\intercal \boldsymbol{y}$ for every point $\boldsymbol{y}$ in P not equal to $\boldsymbol{x}$. If $\boldsymbol{y}$ and $\boldsymbol{z}$ are any two such points and $t \in (0,1)$, then $\boldsymbol{c}^\intercal \boldsymbol{x} = \boldsymbol{c}^\intercal (t\boldsymbol{x} + (1-t)\boldsymbol{x}) < \boldsymbol{c}^\intercal (t\boldsymbol{y} + (1-t)\boldsymbol{z})$. Hence $\boldsymbol{x} \neq t\boldsymbol{y} + (1-t)\boldsymbol{z}$. That is, $\boldsymbol{x}$ cannot be expressed as a convex combination of any two other points of P; it is therefore an extreme point.

Extreme point $\Rightarrow$ Basic feasible solution:
Let $\boldsymbol{x} \in P$ and $I = \{i : \boldsymbol{a}_i^\intercal \boldsymbol{x} = b_i\}$. If $\boldsymbol{x}$ were not a basic feasible solution, there would not exist n linearly independent vectors in the set $\{\boldsymbol{a}_i : i \in I\}$, which would therefore span a proper subset of $\mathbb{R}^n$. Choose a nonzero vector $\boldsymbol{d} \in \mathbb{R}^n$ orthogonal to each of the vectors in this set, so that $\boldsymbol{a}_i^\intercal \boldsymbol{d} = 0$ for each $i \in I$. For $\epsilon > 0$ consider the points $\boldsymbol{y} = \boldsymbol{x} + \epsilon \boldsymbol{d}$ and $\boldsymbol{z} = \boldsymbol{x} - \epsilon \boldsymbol{d}$. We see that $\boldsymbol{a}_i^\intercal \boldsymbol{y} = \boldsymbol{a}_i^\intercal \boldsymbol{z} = b_i$ for each $i \in I$. Moreover, for $j \notin I$, we know that $\boldsymbol{a}_j^\intercal \boldsymbol{x} > b_j$. If we choose ϵ so that $\epsilon \left|\boldsymbol{a}_j^\intercal \boldsymbol{d}\right| < \boldsymbol{a}_j^\intercal \boldsymbol{x} - b_j$ for all $j \notin I$, then $\boldsymbol{a}_j^\intercal \boldsymbol{y} = \boldsymbol{a}_j^\intercal \boldsymbol{x} + \epsilon \boldsymbol{a}_j^\intercal \boldsymbol{d} > b_j$ and likewise $\boldsymbol{a}_j^\intercal \boldsymbol{z} = \boldsymbol{a}_j^\intercal \boldsymbol{x} - \epsilon \boldsymbol{a}_j^\intercal \boldsymbol{d} > b_j$ for each $j \notin I$. Thus, $\boldsymbol{y}$ and $\boldsymbol{z}$ also belong to P. But then $\boldsymbol{x} = (\boldsymbol{y} + \boldsymbol{z})/2$ could not be an extreme point of P.

Basic feasible solution $\Rightarrow$ Vertex:
Let $\boldsymbol{x}$ be a basic feasible solution of P, $I = \{i : \boldsymbol{a}_i^\intercal \boldsymbol{x} = b_i\}$, and $\boldsymbol{c} = \sum_{i \in I} \boldsymbol{a}_i$. Then

$$\boldsymbol{c}^\intercal \boldsymbol{x} = \sum_{i \in I} \boldsymbol{a}_i^\intercal \boldsymbol{x} = \sum_{i \in I} b_i.$$

For any $\boldsymbol{y} \in P$ we know that $\boldsymbol{a}_i^\intercal \boldsymbol{y} \geq b_i$, so that

$$\boldsymbol{c}^\intercal \boldsymbol{y} = \sum_{i \in I} \boldsymbol{a}_i^\intercal \boldsymbol{y} \geq \sum_{i \in I} b_i = \boldsymbol{c}^\intercal \boldsymbol{x}. \tag{2.4}$$

That is, $\boldsymbol{x}$ is an optimal solution to the problem of minimizing $\boldsymbol{c}^\intercal \boldsymbol{y}$ over P. Equality in (2.4) holds iff $\boldsymbol{a}_i^\intercal \boldsymbol{y} = b_i$ for all $i \in I$. But this system of equations has a unique solution since $\boldsymbol{x}$ is a basic feasible solution, with n linearly independent constraints active at $\boldsymbol{x}$. That is, equality holds in (2.4) only at the unique minimizer $\boldsymbol{x}$. This means that $\boldsymbol{x}$ is a vertex of P.

Corollary 2.5.1: Given a finite number of linear constraints, there can only be a finite number of basic solutions.

Proof: Suppose that $k \geq n$ constraints are imposed on a basic solution $\boldsymbol{x} \in \mathbb{R}^n$. We know that n linearly independent constraints must be active at $\boldsymbol{x}$, uniquely defining a point in $\mathbb{R}^n$. Different basic solutions correspond to different sets of n linearly independent active constraints chosen from the k imposed constraints. Therefore, an upper bound to the number of basic solutions is $\binom{k}{n}$.

Remark: The number of basic feasible solutions, while finite, can be very large: the unit hypercube $\{\boldsymbol{x} \in \mathbb{R}^n : 0 \leq x_j \leq 1, j = 1, \ldots, n\}$ has $2n$ constraints but 2^n basic feasible solutions.

Q. What about a half space in $\mathbb{R}^2$? How many basic solutions are there?

Definition: Two distinct basic solutions to a set of linear constraints in $\mathbb{R}^n$ are *adjacent* if they share $n-1$ linearly independent active constraints.

- In Figure 2.2, A and C are adjacent to D; B and D are adjacent to E.

Definition: The line segment joining two adjacent basic feasible solutions is called an *edge*.

2.H. Polyhedra in standard form

To find the basic solutions for a given convex polyhedron, it is convenient to express it in the standard form $\{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq \boldsymbol{0}\}$, where without loss of generality $\boldsymbol{A}$ is an $m \times n$ matrix with linearly independent rows (which requires $m \leq n$).

Remark: Every basic solution must satisfy the m linear independent equality constraints $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$; this provides us with m active constraints. To obtain a total of n active constraints, we need to choose $n - m$ of the n decision variables to be zero such that the resulting set of n active constraints is linearly independent. The following theorem gives us some insight as to how this can be accomplished.

Theorem 2.6: *Consider the constraints $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ and $\boldsymbol{x} \geq \boldsymbol{0}$, where the $m \times n$ matrix $\boldsymbol{A}$ has linearly independent rows. A vector $\boldsymbol{x} \in \mathbb{R}^n$ is a basic solution iff $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ and there exist indices $j_1, \ldots, j_m$ such that*

(i) *the columns $\boldsymbol{A}_{j_1}, \ldots, \boldsymbol{A}_{j_m}$ are linearly independent;*

(ii) *$x_i = 0$ for all $i \neq j_1, \ldots, j_m$.*

Proof: Suppose that $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ and conditions (i) and (ii) hold at some $\boldsymbol{x} \in \mathbb{R}^n$. Then

$$\sum_{i=1}^{m} \boldsymbol{A}_{j_i} x_{j_i} = \sum_{i=1}^{n} \boldsymbol{A}_i x_i = \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}.$$

The linear independence of the columns $\boldsymbol{A}_{j_1}, \ldots, \boldsymbol{A}_{j_m}$ implies that the solution $x_{j_1}, \ldots, x_{j_m}$ of this system of m equations is unique. Since the remaining $n-m$ decision variables are zero, we see that the solution $\boldsymbol{x} = (x_1, \ldots, x_n)$ in $\mathbb{R}^n$ satisfies n linearly independent active constraints. Furthermore, all m equality constraints are active. Thus, $\boldsymbol{x}$ is a basic solution.

Conversely, suppose that $\boldsymbol{x}$ is a basic solution. Denote the nonzero components of $\boldsymbol{x}$ by $x_{j_1}, \ldots, x_{j_k}$. Since $\boldsymbol{x}$ is a basic solution, the system of equations formed by the active constraints $\sum_{i=1}^n \boldsymbol{A}_i x_i = \sum_{i=1}^k \boldsymbol{A}_{j_i} x_{j_i} = \boldsymbol{b}$ and $x_i = 0$, $i \neq j_1, \ldots, j_k$ has a unique solution. If the columns $\boldsymbol{A}_{j_1}, \ldots, \boldsymbol{A}_{j_k}$ were linearly dependent, we could find scalars λ_i (not all zero) such that $\sum_{i=1}^k \boldsymbol{A}_{j_i} \lambda_i = \mathbf{0}$, so that $\sum_{i=1}^k \boldsymbol{A}_{j_i}(x_{j_i} + \lambda_i) = \boldsymbol{b}$, contradicting the uniqueness of the solution. Thus the columns $\boldsymbol{A}_{j_1}, \ldots, \boldsymbol{A}_{j_k}$ are linearly independent, which implies that $k \leq m$. But since $\boldsymbol{A}$ has m linearly independent rows, it must have a total of m linearly independent columns. Hence there exist $m-k$ additional columns $\boldsymbol{A}_{j_{k+1}}, \ldots, \boldsymbol{A}_{j_m}$ of $\boldsymbol{A}$ such that the m columns $\boldsymbol{A}_{j_1}, \ldots, \boldsymbol{A}_{j_m}$ are linearly independent. Since the decision variables corresponding to the other columns of $\boldsymbol{A}$ are zero, we see that conditions (i) and (ii) are both satisfied.

Remark: The previous theorem suggests the following procedure for finding all basic solutions.

Procedure for constructing basic solutions

1. Choose m linearly independent columns $\boldsymbol{A}_{j_1}, \ldots, \boldsymbol{A}_{j_m}$.
2. Let $x_i = 0$ for all $i \neq j_1, \ldots, j_m$.
3. Solve the system of m equations $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ for the unknowns $x_{j_1}, \ldots, x_{j_m}$.

If a basic solution constructed according to this procedure is **nonnegative**, then it is a basic **feasible** solution. Every basic feasible solution is a basic solution, and can thus be obtained from this procedure.

Definition: If $\boldsymbol{x}$ is a basic solution, the variables $x_{j_1}, \ldots, x_{j_m}$ are called *basic variables*; the remaining variables are called *nonbasic*. The columns $\boldsymbol{A}_{j_1}, \ldots, \boldsymbol{A}_{j_m}$ are called the *basic columns* and, since they are linearly independent, they form a *basis* of $\mathbb{R}^m$. We will sometimes refer to bases being *distinct*: distinct bases involve different sets $\{j_1, \ldots, j_m\}$ of *basic indices*; if two bases involve the same set of indices in a different order, they will be viewed as equivalent.

By arranging the m basic columns next to each other, we obtain an $m \times m$ matrix $\boldsymbol{B}$ called a *basis matrix*, which is invertible. We also define a vector $\boldsymbol{x}_B$ composed of the m unknown basic variables:

$$\boldsymbol{B} = [\,\boldsymbol{A}_{j_1}, \ldots, \boldsymbol{A}_{j_m}\,], \qquad \boldsymbol{x}_B = \begin{bmatrix} x_{j_1} \\ \vdots \\ x_{j_m} \end{bmatrix}$$

The basic variables $x_{j_1}, \ldots, x_{j_m}$ are uniquely determined by solving the equation $\boldsymbol{B}\boldsymbol{x}_B = \boldsymbol{b}$:

$$\boldsymbol{x}_B = \boldsymbol{B}^{-1}\boldsymbol{b}.$$

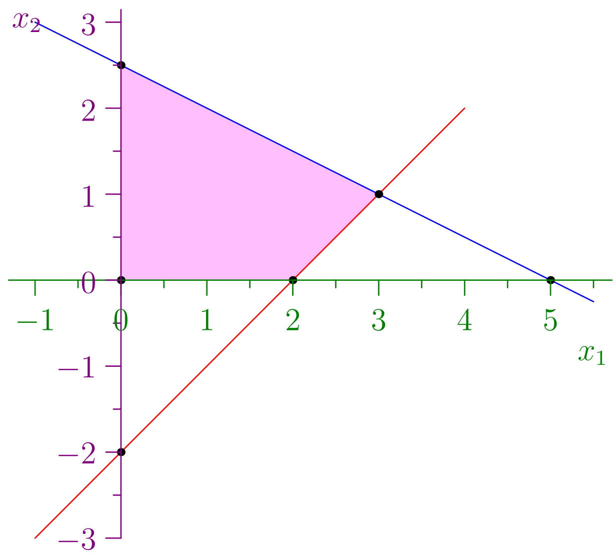


Figure 2.3: Determining basic solutions.

Problem 2.3: Find the vertices of the polyhedron P defined by

$$\begin{aligned} x_1 - x_2 &\le 2, \\ x_1 + 2x_2 &\le 5, \\ x_1, x_2 &\ge 0, \end{aligned}$$

illustrated in Figure 2.3.

To put this problem in standard form, we need to add two slack variables x_3 and x_4, yielding a total of $n = 4$ decision variables satisfying $m = 2$ equality constraints:

$$\boldsymbol{A} = \begin{bmatrix} 1 & -1 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{bmatrix}, \qquad \boldsymbol{b} = \begin{bmatrix} 2 \\ 5 \end{bmatrix}.$$

Now we consider all possible choices for the basis matrix $\boldsymbol{B}$, corresponding to all possible ways of choosing two linearly independent column vectors from $\boldsymbol{A}$.

For $j_1 = 1$ and $j_2 = 2$:

$$\boldsymbol{B} = \begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix}, \qquad \boldsymbol{x}_B = \begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 2 \\ 5 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}.$$

This corresponds to the basic solution $\boldsymbol{x} = (3, 1, 0, 0)$. Since all elements of $\boldsymbol{x}$ are nonnegative, we see that this is actually a basic feasible solution, corresponding to the vertex $(3, 1)$ in Figure 2.3.

For $j_1 = 1$ and $j_2 = 3$:

$$\boldsymbol{B} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \qquad \boldsymbol{x}_B = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 2 \\ 5 \end{bmatrix} = \frac{1}{-1} \begin{bmatrix} 0 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 5 \\ -3 \end{bmatrix}.$$

This corresponds to the basic solution $\boldsymbol{x} = (5, 0, -3, 0)$. Since the third element of $\boldsymbol{x}$ is negative, we see that this basic solution, corresponding to the dot at $(5, 0)$ in Figure 2.3, is infeasible.

For $j_1 = 1$ and $j_2 = 4$:

$$\boldsymbol{B} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \qquad \boldsymbol{x}_B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 2 \\ 5 \end{bmatrix} = \frac{1}{1} \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}.$$

This corresponds to the basic solution $\boldsymbol{x} = (2, 0, 0, 3)$. Since all elements of $\boldsymbol{x}$ are nonnegative, we see that this is actually a basic feasible solution, corresponding to the vertex $(2, 0)$ in Figure 2.3.

For $j_1 = 2$ and $j_2 = 3$:

$$\boldsymbol{B} = \begin{bmatrix} -1 & 1 \\ 2 & 0 \end{bmatrix}, \qquad \boldsymbol{x}_B = \begin{bmatrix} -1 & 1 \\ 2 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 2 \\ 5 \end{bmatrix} = \frac{1}{-2} \begin{bmatrix} 0 & -1 \\ -2 & -1 \end{bmatrix} \begin{bmatrix} 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 5/2 \\ 9/2 \end{bmatrix}.$$

This corresponds to the basic solution $\boldsymbol{x} = (0, 5/2, 9/2, 0)$. Since all elements of $\boldsymbol{x}$ are nonnegative, we see that this is actually a basic feasible solution, corresponding to the vertex $(0, 5/2)$ in Figure 2.3.

For $j_1 = 2$ and $j_2 = 4$:

$$\boldsymbol{B} = \begin{bmatrix} -1 & 0 \\ 2 & 1 \end{bmatrix}, \qquad \boldsymbol{x}_B = \begin{bmatrix} -1 & 0 \\ 2 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 2 \\ 5 \end{bmatrix} = \frac{1}{-1} \begin{bmatrix} 1 & 0 \\ -2 & -1 \end{bmatrix} \begin{bmatrix} 2 \\ 5 \end{bmatrix} = \begin{bmatrix} -2 \\ 9 \end{bmatrix}.$$

This corresponds to the basic solution $\boldsymbol{x} = (0, -2, 0, 9)$. Since the second element of $\boldsymbol{x}$ is negative, we see that this basic solution, corresponding to the dot at $(0, -2)$ in Figure 2.3, is infeasible.

For $j_1 = 3$ and $j_2 = 4$:

$$\boldsymbol{B} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad \boldsymbol{x}_B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 2 \\ 5 \end{bmatrix} = \frac{1}{1} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \end{bmatrix}.$$

This corresponds to the basic solution $\boldsymbol{x} = (0, 0, 2, 5)$. Since all elements of $\boldsymbol{x}$ are nonnegative, we see that this is actually a basic feasible solution, corresponding to the vertex $(0, 0)$ in Figure 2.3.

We thus see that our procedure has found all $\binom{n}{m} = \binom{4}{2} = 6$ basic solutions, and determined that four of them are feasible, precisely as observed in Figure 2.3.

2.I. Correspondence of bases and basic solutions

Different basic solutions correspond to different bases. However, two different bases may lead to the same basic solutions, e.g., when $\boldsymbol{b} = \mathbf{0}$.

Definition: Two basis matrices are *adjacent* if they share all but one basic column. Adjacent basic solutions can always be obtained from two adjacent bases. Conversely, if two adjacent bases lead to distinct basic solutions, the latter are adjacent.

- In Prob. 2.3, we see that adjacent basis matrices produce adjacent basic solutions.

2.J. Degeneracy

Definition: A basic solution $\boldsymbol{x} \in \mathbb{R}^n$ is *degenerate* if more than n constraints are active at $\boldsymbol{x}$.

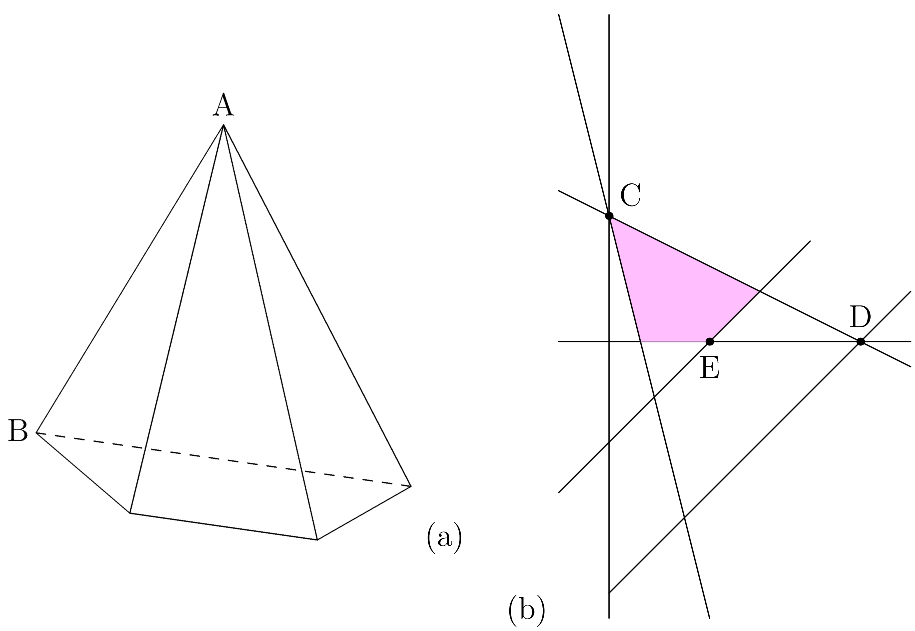


Figure 2.4: (a) a convex polyhedron in $\mathbb{R}^3$; (b) a convex polyhedron in $\mathbb{R}^2$.

- Which of the basic solutions illustrated in Figure 2.4 are degenerate and which are feasible?
  A: degenerate basic feasible
  B: nondegenerate basic feasible
  C: degenerate basic feasible
  D: degenerate basic infeasible
  E: nondegenerate basic feasible

Remark: Consider the standard-form polyhedron $P = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq \boldsymbol{0}\}$, where $\boldsymbol{A}$ is an $m \times n$ matrix. Let $\boldsymbol{x}$ be a basic solution. The vector $\boldsymbol{x}$ is a degenerate basic solution if **more than** $n - m$ of the components of $\boldsymbol{x}$ are zero.

Remark: Since in standard form the $n - m$ nonbasic variables of a basic solution must be zero, **a basic solution is degenerate iff at least one of the basic variables is zero.**

Remark: Degeneracy is not a purely geometric property. Consider the polyhedron

$$P = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} : x_1 - x_2 = 0, x_1 + x_2 + 2x_3 = 2, x_1, x_2, x_3 \geq 0 \right\}.$$

Which of the following basic solutions are degenerate? Here $n = 3$ and $m = 2$.

$\begin{pmatrix} 1 & 1 & 0 \end{pmatrix}^\intercal$: nondegenerate (3 active constraints)
$\begin{pmatrix} 0 & 0 & 1 \end{pmatrix}^\intercal$: degenerate (4 active constraints)

Now consider the polyhedron represented by the same set but with the constraint $x_2 \geq 0$ relaxed:

$$P = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} : x_1 - x_2 = 0, x_1 + x_2 + 2x_3 = 2, x_1, x_3 \geq 0 \right\}.$$

Now which of the following are degenerate? Here $n = 3$ and $m = 2$.

$\begin{pmatrix} 1 & 1 & 0 \end{pmatrix}^\intercal$: nondegenerate (3 active constraints)
$\begin{pmatrix} 0 & 0 & 1 \end{pmatrix}^\intercal$: nondegenerate (3 active constraints)

Remark: For standard-form representations, if a basic solution is degenerate, it is degenerate under every standard-form representation.

Remark: We construct basic solutions by choosing n linearly independent constraints to be satisfied with equality. Degeneracy arises when additional constraints are also satisfied with equality at that basic solution.

2.K. Existence of extreme points

Definition: A polyhedron $P \subset \mathbb{R}^n$ *contains a line* if there exists a vector $\boldsymbol{x} \in P$ and a nonzero vector $\boldsymbol{d} \in \mathbb{R}^n$ such that $\boldsymbol{x} + t\boldsymbol{d} \in P$ for all scalars $t \in \mathbb{R}$.

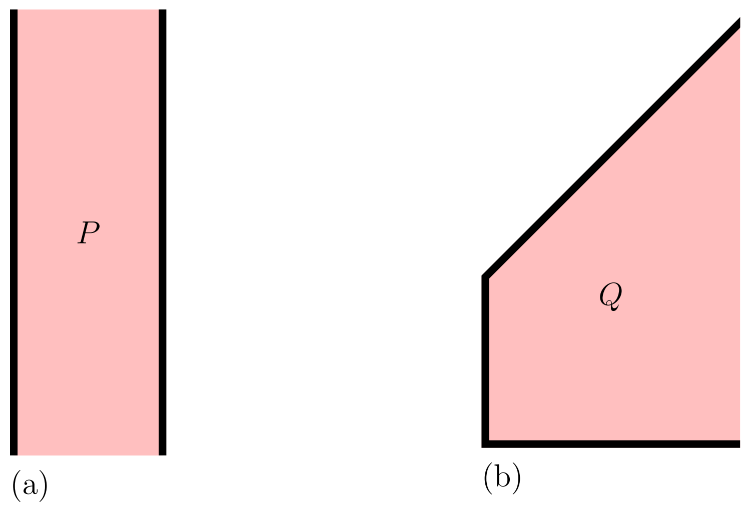


Figure 2.5: (a) P contains a line; (b) Q does not contain a line.

Remark: A bounded polyhedron does not contain a line.

Remark: A convex polyhedron in standard form is contained in the positive orthant and hence does not contain a line.

Remark: Notice for $n > 1$ that a half space in $\mathbb{R}^n$ contains a line, but has no extreme points. The following theorem states that these two properties are equivalent.

Theorem 2.7: *Suppose that the polyhedron* $P = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{a}_i^\intercal \boldsymbol{x} \geq b_i, i = 1, \ldots, k\}$ *is nonempty. Then the following are equivalent:*

(i) *P does not contain a line.*

(ii) *P has at least one extreme point.*

(iii) *The set $\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k\}$ contains n linearly independent vectors.*

Proof:

(i) $\Rightarrow$ (ii): Let $\boldsymbol{x}$ be an element of P and let $I = \{i : \boldsymbol{a}_i^\mathsf{T}\boldsymbol{x} = b_i\}$. If n of the vectors $\boldsymbol{a}_i$, $i \in I$, are linearly independent then $\boldsymbol{x}$ is a basic feasible solution (extreme point). Otherwise, these vectors span a proper subspace of $\mathbb{R}^n$. Let $\boldsymbol{d} \neq \boldsymbol{0}$ be a vector orthogonal to this subspace, so that $\boldsymbol{a}_i^\mathsf{T}\boldsymbol{d} = 0$ for every $i \in I$. Consider $\boldsymbol{y} = \boldsymbol{x} + t\boldsymbol{d}$ for $t \in \mathbb{R}$. For $i \in I$, we know that $\boldsymbol{a}_i^\mathsf{T}\boldsymbol{y} = \boldsymbol{a}_i^\mathsf{T}\boldsymbol{x} + t\boldsymbol{a}_i^\mathsf{T}\boldsymbol{d} = \boldsymbol{a}_i^\mathsf{T}\boldsymbol{x} = b_i$. That is, all constraints active at $\boldsymbol{x}$ remain active on the line $L = \{\boldsymbol{x} + t\boldsymbol{d} : t \in \mathbb{R}\}$. However, since P contains no lines, its boundary must have an intersection with L at some point $\boldsymbol{x} + T\boldsymbol{d}$, where a new constraint, say $\boldsymbol{a}_j^\mathsf{T}(\boldsymbol{x} + T\boldsymbol{d}) = b_j$ for some $j \notin I$, becomes active. Now $j \notin I$ implies that $\boldsymbol{a}_j^\mathsf{T}\boldsymbol{x} \neq b_j$, so $\boldsymbol{a}_j^\mathsf{T}\boldsymbol{d} \neq 0$ and hence $\boldsymbol{a}_j$ is not a linear combination of $\boldsymbol{a}_i$, $i \in I$. Thus, by moving from $\boldsymbol{x}$ to $\boldsymbol{x} + T\boldsymbol{d}$, we have increased the number of linearly independent active constraints by at least one. Using induction, we then see on repeating this argument that we will eventually reach a basic solution where n linearly independent constraints are active. Since these movements are always confined within P, such a point is a basic feasible solution (extreme point).

(ii) $\Rightarrow$ (iii): If P has an extreme point $\boldsymbol{x}$, then by definition (of a basic feasible solution) there exists n linearly independent active constraints at $\boldsymbol{x}$, corresponding to n linearly independent active constraint vectors.

(iii) $\Rightarrow$ (i): Suppose n of the vectors $\boldsymbol{a}_i$ are linearly independent. Let us relabel them $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$. If P were to contain a line $\{\boldsymbol{x} + t\boldsymbol{d} : t \in \mathbb{R}\}$, where $\boldsymbol{d} \neq \boldsymbol{0}$, then $\boldsymbol{a}_i^\mathsf{T}(\boldsymbol{x} + t\boldsymbol{d}) \geq b_i$ for all i and all $t \in \mathbb{R}$. The only way that can happen is if $\boldsymbol{a}_i^\mathsf{T}\boldsymbol{d} = 0$ for all i. The linear independence of the vectors $\boldsymbol{a}_i$ then implies that $\boldsymbol{d} = \boldsymbol{0}$. This contradiction establishes that P cannot contain a line.

Corollary 2.7.1: Every nonempty bounded convex polyhedron has at least one basic feasible solution.

Corollary 2.7.2: Every nonempty convex polyhedron in standard form has at least one basic feasible solution.

2.L. Optimality of extreme points

Theorem 2.8: *Consider the linear programming problem of minimizing $\boldsymbol{c}^\mathsf{T}\boldsymbol{x}$ over a convex polyhedron P. Suppose that P has at least one extreme point and that there exists an optimal solution. Then there exists an optimal solution that is an extreme point of P.*

Proof: Express $P = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b}\}$. Let v be the minimal value of the cost $\boldsymbol{c}^\mathsf{T}\boldsymbol{x}$. Consider the polyhedron $Q = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b}, \boldsymbol{c}^\mathsf{T}\boldsymbol{x} = v\}$, the (nonempty) set of all optimal solutions. By Theorem 2.7, we see that P contains no lines, and hence $Q \subset P$ also contains no lines and therefore has an extreme point, say $\boldsymbol{x}^*$. If $\boldsymbol{x}^*$ were not also an extreme point of P, there would exists points $\boldsymbol{y}, \boldsymbol{z} \in P$ distinct from $\boldsymbol{x}^*$ and some $t \in (0, 1)$ such that $\boldsymbol{x}^* = t\boldsymbol{y} + (1 - t)\boldsymbol{z}$ and hence

$v = \boldsymbol{c}^\intercal \boldsymbol{x}^* = t\boldsymbol{c}^\intercal \boldsymbol{y} + (1-t)\boldsymbol{c}^\intercal \boldsymbol{z} \geq tv + (1-t)v = v$ since v is the optimal cost. Then $\boldsymbol{c}^\intercal \boldsymbol{y} = \boldsymbol{c}^\intercal \boldsymbol{z} = v$, so that $\boldsymbol{y}$ and $\boldsymbol{z}$ would both belong to Q. But this would contradict the fact that $\boldsymbol{x}^*$ is an extreme point of Q. Thus the optimal solution $\boldsymbol{x}^*$ is also an extreme point of P.

Remark: By keeping track of the costs in Theorem 2.7 as we move towards an extreme point, one can prove a strengthened version of Theorem 2.8, which shows that if the optimal cost is finite, the existence of an optimal solution is guaranteed:

Theorem 2.9: *Consider the linear programming problem of minimizing $\boldsymbol{c}^\intercal \boldsymbol{x}$ over a convex polyhedron P. Suppose that P has at least one extreme point. Then either the optimal cost equals $-\infty$ or P has an optimal extreme point.*

Remark: Since every linear programming problem can be transformed into an equivalent standard-form problem, and every nonempty convex polyhedron in standard form has at least one extreme point (Corollary 2.7.2), we can further simplify this result to:

Corollary 2.9.1: Consider the linear programming problem of minimizing $\boldsymbol{c}^\intercal \boldsymbol{x}$ over a nonempty convex polyhedron P. Then either the optimal cost equals $-\infty$ or there exists an optimal solution at an extreme point of P.

Remark: The previous result does not hold for nonlinear cost functions. Consider the problem of minimizing $1/x$ subject to $x \geq 1$: the optimal cost is 0 but an optimal solution does not exist!

Chapter 3

The Simplex Method

In this chapter we develop a general method, called the *simplex method* for solving linear programming problems written in standard form. A *simplex* extends the notion of a line segment in $\mathbb{R}^1$, a triangle in $\mathbb{R}^2$, or a tetrahedron in $\mathbb{R}^3$ to arbitrary dimensions:

Definition: A n-*simplex* is a convex polyhedron in $\mathbb{R}^n$ that is the convex hull of its $n+1$ vertices $\boldsymbol{v}_i \in \mathbb{R}^n$, $i = 0, \ldots, n$:

$$\left\{ t_0\boldsymbol{v}_0 + \ldots + t_n\boldsymbol{v}_n : \sum_{i=0}^{n} t_i = 1, \quad t_i \geq 0, \quad i = 0\ldots, n \right\}.$$

Definition: The *standard* n-*simplex* in $\mathbb{R}^{n+1}$ is the n-simplex obtained when the vectors $\boldsymbol{v}_i$ are chosen to be the unit vectors $\boldsymbol{e}_i$:

$$\left\{ (t_0, \ldots, t_n) \in \mathbb{R}^{n+1} : \sum_{i=0}^{n} t_i = 1, \quad t_i \geq 0, \quad i = 0\ldots, n \right\}.$$

The standard n-simplex always lies in the first orthant.

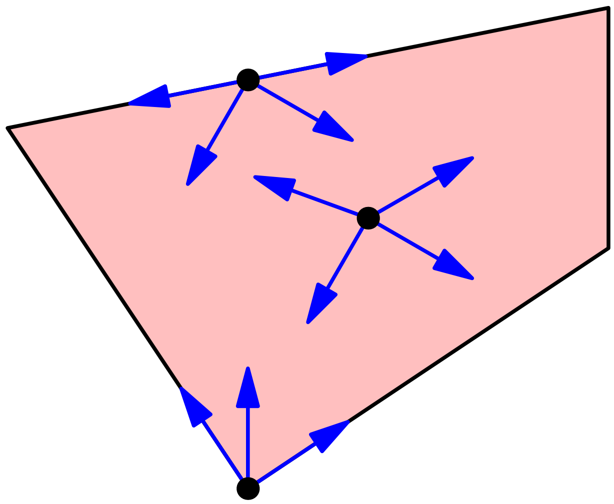
We will develop the simplex method for the standard-form problem

$$\begin{aligned} \text{minimize} \quad & \boldsymbol{c}^\mathsf{T}\boldsymbol{x} \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \\ & \boldsymbol{x} \geq \boldsymbol{0}. \end{aligned}$$

Let P be the corresponding feasible set and $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ have m linearly independent rows. We will continue to denote the ith row of $\boldsymbol{A}$ as $\boldsymbol{a}_i$ and the jth column of $\boldsymbol{A}$ as $\boldsymbol{A}_j$.

Definition: Let $\boldsymbol{x}$ be an element of the convex polyhedron P. A nonzero vector $\boldsymbol{d} \in \mathbb{R}^n$ is a *feasible direction* at $\boldsymbol{x}$ if there exists a positive scalar t for which $\boldsymbol{x} + t\boldsymbol{d} \in P$.

- Which of the directions in the following figure are feasible?



Remark: Let $\boldsymbol{x} \in \mathbb{R}^n$ be a basic feasible solution for the linear programming problem $\boldsymbol{Ax} = \boldsymbol{b}$, $\boldsymbol{x} \geq \boldsymbol{0}$ associated with basis matrix $\boldsymbol{B}$ composed of columns $\{j_1, \ldots, j_m\}$ of $\boldsymbol{A}$. Suppose we move from $\boldsymbol{x}$ to $\boldsymbol{x} + t\boldsymbol{d}$ by selecting a nonbasic variable x_j (which is initially 0 since $j \notin \{j_1, \ldots, j_m\}$) and increasing it by $t > 0$. That is, we choose $d_j = 1$ and set $d_i = 0$ for all other nonbasic indices $i \neq j$. The remaining (basic) components of $\boldsymbol{d}$ still need to be determined. Since we are only interested in feasible solutions, we require $\boldsymbol{A}(\boldsymbol{x} + t\boldsymbol{d}) = \boldsymbol{b}$. Thus,

$$\left.\begin{array}{l} \boldsymbol{A}(\boldsymbol{x} + t\boldsymbol{d}) = \boldsymbol{b} \\ \boldsymbol{Ax} = \boldsymbol{b} \end{array}\right\} \Rightarrow \boldsymbol{Ad} = \boldsymbol{0} \Rightarrow \sum_{i=1}^{n} \boldsymbol{A}_i d_i = \sum_{i=1}^{m} \boldsymbol{A}_{j_i} d_{j_i} + \boldsymbol{A}_j = \boldsymbol{B}\boldsymbol{d}_B + \boldsymbol{A}_j = \boldsymbol{0},$$

where

$$\boldsymbol{d}_B = -\boldsymbol{B}^{-1}\boldsymbol{A}_j$$

contains the basic components $(d_{j_1}, \ldots, d_{j_m})$ of $\boldsymbol{d}$.

Definition: The direction vector $\boldsymbol{d}$ constructed above is called the jth *simplex direction*.

So far, the equality constraints are guaranteed to be satisfied along this direction. How about the inequality constraints: $\boldsymbol{x} \geq \boldsymbol{0}$?

Q. Are they respected for nonbasic variables?

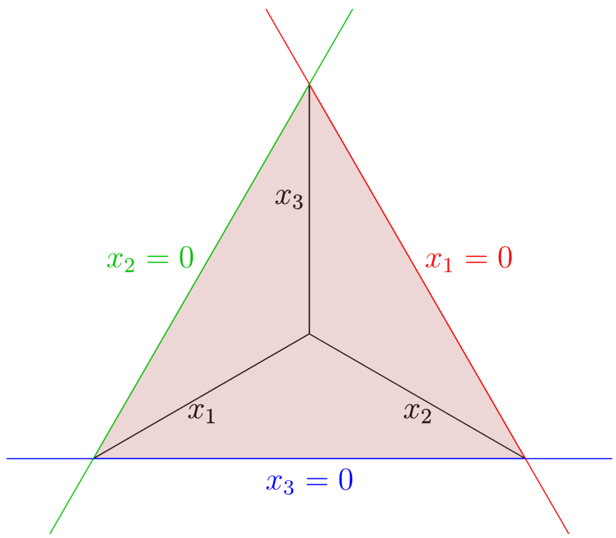
A. Yes, since the nonbasic variables are initially zero and only the jth one is adjusted, by a positive amount t.

Remark: For the basic variables, we distinguish the following cases:

(i) $\boldsymbol{x}$ is a nondegenerate basic feasible solution. Then $\boldsymbol{x}_B > \mathbf{0}$, yielding $\boldsymbol{x}_B + t\boldsymbol{d}_B > \mathbf{0}$ for sufficiently small t. Thus $\boldsymbol{d}$ is a feasible direction.

(ii) $\boldsymbol{x}$ is a degenerate basic feasible solution. Then $\boldsymbol{d}$ is not always a feasible direction. It is possible that a basic variable x_{j_i} is zero and d_{j_i} is negative.

- Let $n = 3, m = 1, n - m = 2$ and $\boldsymbol{A} = [1, 1, 1]$, $\boldsymbol{b} = 1$. Let us visualize the feasible set on the two-dimensional plane $x_1 + x_2 + x_3 = 1$, where its edges are associated with the nonnegativity constraints $[x_1, x_2, x_3] \geq \mathbf{0}$.

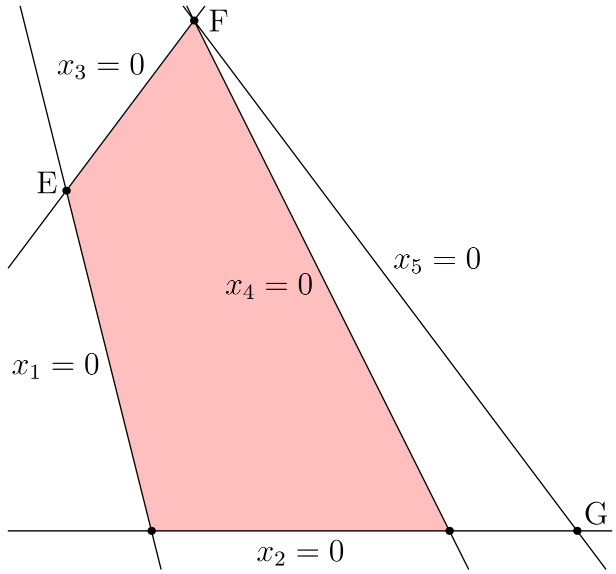


- Let $n = 5, m = 3, n - m = 2$. In the following figure, we can visualize the feasible region by focusing on the two-dimensional set defined by the constraint $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$. The edges of the feasible set are associated with the nonnegativity constraints $\boldsymbol{x} \geq \mathbf{0}$.

Q. What kind of solution occurs at point E?
nondegenerate basic feasible solution

Q. What are the basic and nonbasic variables at E?
nonbasic: x_1 and x_3 basic: x_2, x_4, x_5

Q. Consider the direction obtained by increasing x_1 while holding x_3 zero. Is this a simplex direction? Yes.



Q. Is it also feasible? Yes.

Q. What kind of solution occurs at point F? Degenerate basic feasible solution.

Q. Give examples of basic and nonbasic variables at F.

nonbasic: x_3 and x_4 basic: x_1, x_2, x_5
nonbasic: x_3 and x_5 basic: x_1, x_2, x_4
nonbasic: x_4 and x_5 basic: x_1, x_2, x_3

Q. Suppose the nonbasic variables are x_3 and x_5. Consider the direction obtained by increasing x_3 while holding the other nonbasic variable $x_5 = 0$ ($d_3 = 1$, $d_5 = 0$). Is this a simplex direction? Yes.

Q. Is it also feasible? No.

Remark: We need to know the effect on the cost when moving by an amount t in the jth simplex direction $\boldsymbol{d}$. The change in the cost is

$$\boldsymbol{c}^\mathsf{T}(\boldsymbol{x} + t\boldsymbol{d}) - \boldsymbol{c}^\mathsf{T}\boldsymbol{x} = t\boldsymbol{c}^\mathsf{T}\boldsymbol{d}.$$

Definition: Let $\boldsymbol{c}_B = (c_{j_1}, \ldots, c_{j_m})$ be the *basic cost vector*.

Remark: The rate of cost change $\boldsymbol{c}^\intercal \boldsymbol{d}$ along the jth simplex direction $\boldsymbol{d}$ is

$$\boldsymbol{c}^\intercal \boldsymbol{d} = \boldsymbol{c}_B^\intercal \boldsymbol{d}_B + c_j = c_j - \boldsymbol{c}_B^\intercal \boldsymbol{B}^{-1} \boldsymbol{A}_j.$$

Here c_j represents the change in the cost per unit increase in the non basic variable variable x_j and the term $\boldsymbol{c}_B^\intercal \boldsymbol{B}^{-1} \boldsymbol{A}_j$ represents the cost of the compensating change in the basic variables required to satisfy the constraint $\boldsymbol{Ax} = \boldsymbol{b}$.

Definition: Let $\boldsymbol{x}$ be a basic solution, $\boldsymbol{B}$ be an associated basis matrix, and $\boldsymbol{c}_B$ be the associated costs of the basic variables. For each j, we define the *reduced cost* $\bar{c}_j$ of the variable x_j to be the rate of change in cost along the jth simplex direction:

$$\bar{c}_j = c_j - \boldsymbol{c}_B^\intercal \boldsymbol{B}^{-1} \boldsymbol{A}_j.$$

Problem 3.1: Consider the linear programming problem

$$\begin{array}{lccccccccc} \text{minimize} & c_1x_1 & + & c_2x_2 & + & c_3x_3 & + & c_4x_4 & & \\ \text{subject to} & x_1 & + & x_2 & + & x_3 & + & x_4 & = & 2, \\ & 2x_1 & + & & + & 3x_3 & + & 4x_4 & = & 2, \\ & \multicolumn{9}{l}{x_1, x_2, x_3, x_4 \geq 0\ .} \end{array}$$

Compute the reduced cost $\bar{c}_3$ of moving in the 3rd simplex direction from the basic solution associated with the basis matrix

$$\boldsymbol{B} = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}.$$

From the constraint matrix

$$\boldsymbol{A} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 0 & 3 & 4 \end{bmatrix},$$

we see that $\boldsymbol{A}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\boldsymbol{A}_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ are linearly independent: they form a basis of $\mathbb{R}^2$. Let $\boldsymbol{b} = (2, 2)$. We now determine the basic variable vector

$$\boldsymbol{x}_B = \boldsymbol{B}^{-1}\boldsymbol{b} = \frac{1}{-2}\begin{bmatrix} 0 & -1 \\ -2 & 1 \end{bmatrix}\begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

which corresponds to the nondegenerate basic feasible solution $\boldsymbol{x} = (1, 1, 0, 0)$. The simplex direction corresponding to increasing the nonbasic variable x_3 has nonbasic components $d_3 = 1$ and $d_4 = 0$ and basic components

$$\begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = \begin{bmatrix} d_{j_1} \\ d_{j_2} \end{bmatrix} = \boldsymbol{d}_B = -\boldsymbol{B}^{-1}\boldsymbol{A}_3 = -\frac{1}{-2}\begin{bmatrix} 0 & -1 \\ -2 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} -3/2 \\ 1/2 \end{bmatrix}.$$

Then the reduced cost of moving in the 3rd simplex direction is

$$\bar{c}_3 = c_3 + \boldsymbol{c}_B^\intercal \boldsymbol{d}_B = c_3 + [\, c_1 \quad c_2 \,]\begin{bmatrix} -3/2 \\ 1/2 \end{bmatrix} = -\frac{3}{2}c_1 + \frac{1}{2}c_2 + c_3.$$

Definition: Let $\boldsymbol{e}_i$ denote the ith unit vector.

Remark: What is the reduced cost of a basic variable? We find

$$\bar{c}_{j_i} = c_{j_i} - \boldsymbol{c}_B^\intercal \boldsymbol{B}^{-1}\boldsymbol{A}_{j_i} = c_{j_i} - \boldsymbol{c}_B^\intercal \boldsymbol{e}_i = c_{j_i} - c_{j_i} = 0.$$

That is, the reduced cost of every basic variable is zero, so we never need to compute the reduced cost of a basic variable!

Definition: The *reduced cost vector* is given by $\bar{\boldsymbol{c}}^\intercal = \boldsymbol{c}^\intercal - \boldsymbol{c}_B^\intercal \boldsymbol{B}^{-1}\boldsymbol{A}$.

Is there a relationship between the reduced cost and an optimal solution?

Theorem 3.1: *Consider a basic feasible solution* $\boldsymbol{x}$ *associated with a basis matrix* $\boldsymbol{B}$, *and let* $\bar{\boldsymbol{c}}$ *be the corresponding reduced cost vector. Then*

(i) *if* $\bar{\boldsymbol{c}} \geq \boldsymbol{0}$, $\boldsymbol{x}$ *is optimal.*

(ii) *if* $\boldsymbol{x}$ *is optimal and nondegenerate, then* $\bar{\boldsymbol{c}} \geq \boldsymbol{0}$.

Proof:

Let $\boldsymbol{y}$ be an arbitrary **feasible** solution. The difference vector $\boldsymbol{d} = \boldsymbol{y} - \boldsymbol{x}$ satisfies $\boldsymbol{A}\boldsymbol{d} = \boldsymbol{A}\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} - \boldsymbol{b} = \boldsymbol{0}$. That is,

$$\boldsymbol{B}\boldsymbol{d}_B + \sum_{j\in N} \boldsymbol{A}_j d_j = \boldsymbol{0},$$

where N is the set of nonbasic indices. Thus

$$\boldsymbol{d}_B = -\sum_{j\in N} \boldsymbol{B}^{-1}\boldsymbol{A}_j d_j.$$

The change in the cost by moving from $\boldsymbol{x}$ to $\boldsymbol{y}$ is therefore

$$\boldsymbol{c}^\intercal\boldsymbol{y} - \boldsymbol{c}^\intercal\boldsymbol{x} = \boldsymbol{c}^\intercal\boldsymbol{d} = \boldsymbol{c}_B^\intercal\boldsymbol{d}_B + \sum_{j\in N} c_j d_j = \sum_{j\in N}(c_j - \boldsymbol{c}_B^\intercal\boldsymbol{B}^{-1}\boldsymbol{A}_j)d_j = \sum_{j\in N}\bar{c}_j d_j.$$

Thus, $\bar{c}_j$ *is the rate of cost change along the* j*th simplex direction.* Note for every $j \in N$ that the nonbasic variable $x_j = 0$ and $y_j \geq 0$, so $d_j \geq 0$ for all $j \in N$.

i) Given $\bar{\boldsymbol{c}} \geq \boldsymbol{0}$, we then see that

$$\boldsymbol{c}^\intercal\boldsymbol{y} - \boldsymbol{c}^\intercal\boldsymbol{x} = \sum_{j\in N}\bar{c}_j d_j \geq 0.$$

That is, $\boldsymbol{x}$ is an optimal basic feasible solution.

ii) Suppose that $\boldsymbol{x}$ is a nondegenerate basic feasible solution and that $\bar{c}_j < 0$ for some j. Since the reduced cost of a basic variable is always zero, x_j must be a nonbasic variable. Since $\boldsymbol{x}$ is nondegenerate and $\bar{c}_j < 0$, the jth simplex direction is a feasible direction of cost decrease and therefore $\boldsymbol{x}$ is not optimal.

Remark: If $\boldsymbol{x}$ is a degenerate optimal basic feasible solution, it is still possible that $\bar{c}_j < 0$ for some nonbasic index j.

Problem 3.2: Let $\boldsymbol{x}$ be a basic feasible solution of a linear programming problem Π written in standard form, with associated basis matrix $\boldsymbol{B}$ and set of nonbasic indices N. Let $\boldsymbol{y}$ be any feasible solution to Π and consider the difference vector $\boldsymbol{d} = \boldsymbol{y} - \boldsymbol{x}$.

(a) Prove that $d_j \geq 0$ for every $j \in N$.

(b) If $d_j = 0$ for every $j \in N$, prove that $\boldsymbol{y} = \boldsymbol{x}$.

(c) If the reduced cost $\bar{c}_j$ of every nonbasic variable x_j is positive, use parts (a) and (b) to prove that $\boldsymbol{x}$ is the unique optimal solution to Π.

(d) Suppose that Π is nondegenerate and that $\boldsymbol{x}$ is an optimal solution to Π. If the reduced cost $\bar{c}_j$ of some nonbasic variable x_j is zero, prove that Π does not have a unique optimal solution.

Definition: A basis matrix $\boldsymbol{B}$ is *optimal* if both of the following conditions are satisfied:

1. feasibility: $\boldsymbol{x}_B \geq \boldsymbol{0}$;
2. nonnegativity of the reduced costs: $\bar{\boldsymbol{c}} \geq \boldsymbol{0}$.

Remark: The basic solution corresponding to an optimal basis matrix is feasible and satisfies the optimality conditions $\bar{\boldsymbol{c}} \geq \boldsymbol{0}$. It is therefore an optimal solution. On the other hand, in the degenerate case, having an optimal basic solution does not necessarily mean that the reduced costs are nonnegative.

3.A. Development of the simplex method

For now, assume that every basic feasible solution is nondegenerate. Suppose that we are at a basic feasible solution $\boldsymbol{x}$, for which we have computed the reduced costs $\bar{c}_j$ of the nonbasic variables. Then

- if all of them are nonnegative, according to Theorem 3.1, we have an optimal solution and we stop;
- if the reduced cost $\bar{c}_j$ of some nonbasic variable x_j is negative, the jth simplex direction $\boldsymbol{d}$ is a feasible direction of cost decrease. This direction is obtained by letting $d_j = 1$, $d_i = 0$ for $i \notin \{j, j_1, \ldots, j_m\}$, and $\boldsymbol{d}_B = -\boldsymbol{B}^{-1}\boldsymbol{A}_j$.

Remark: In moving along this direction $\boldsymbol{d}$, the nonbasic variable $\boldsymbol{x}_j$ becomes positive, while all other nonbasic variables are fixed at zero. We describe this situation by saying that x_j (or $\boldsymbol{A}_j$) *enters* or *is brought into the basis*.

Q. How far should we move along the direction $\boldsymbol{d}$?

A. We want to move as far as possible in this direction, while staying inside the polyhedron.

Remark: Suppose we move from $\boldsymbol{x} \to \boldsymbol{x} + t\boldsymbol{d}$. The maximum possible value of t is given by

$$t^* = \max\{t \geq 0 : \boldsymbol{x} + t\boldsymbol{d} \in P\}.$$

Can we provide a formula for t^*? The only condition on t is to keep $\boldsymbol{x}+t\boldsymbol{d}$ feasible. Since $\boldsymbol{A}\boldsymbol{d} = \boldsymbol{0}$, we know that $\boldsymbol{A}(\boldsymbol{x} + t\boldsymbol{d}) = \boldsymbol{A}\boldsymbol{x} + t\boldsymbol{A}\boldsymbol{d} = \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ is always satisfied. That is, the equality constraints are never violated. It remains to check whether the entries of $\boldsymbol{x} + t\boldsymbol{d}$ are nonnegative. There are two cases:

1. If $\boldsymbol{d} \geq \boldsymbol{0}$, then $\boldsymbol{x} + t\boldsymbol{d} \geq \boldsymbol{0}$ for all $t \geq 0$. The vector $\boldsymbol{x} + t\boldsymbol{d}$ never becomes infeasible, so $t^* = \infty$.

2. If $d_i < 0$ for some i, the constraint $x_i + td_i \geq 0$ becomes $t \leq -x_i/d_i$. This condition must hold for every i such that $d_i < 0$. For each nonbasic variable x_i, d_i is nonnegative: either $d_i = 1$ or $d_i = 0$. We therefore only need to consider the basic components $\{d_{j_i} : i = 1, \ldots, m\}$ of $\boldsymbol{d}$:

$$t^* = \min_{\substack{i=1,\ldots,m: \\ d_{j_i}<0}} \left(-\frac{x_{j_i}}{d_{j_i}} \right).$$

   Note that $t^* \geq 0$ since $x_{j_i} \geq 0$ at a basic feasible solution. At a nondegenerate basic feasible solution one has the stronger result that $x_{j_i} > 0$ and hence $t^* > 0$.

Remark: If a finite and positive t^* is found, one can move to the new feasible solution $\boldsymbol{y} = \boldsymbol{x} + t^*\boldsymbol{d}$. Since $x_j = 0$ and $d_j = 1$, we have $y_j = t^* > 0$. Furthermore, on letting ℓ be a minimizing index:

$$t^* = -\frac{x_{j_\ell}}{d_{j_\ell}} = \min_{\substack{i=1,\ldots,m: \\ d_{j_i}<0}} \left(-\frac{x_{j_i}}{d_{j_i}} \right)$$

we see that

$$d_{j_\ell} < 0 \qquad \text{and} \qquad y_{j_\ell} = x_{j_\ell} + t^* d_{j_\ell} = 0.$$

We observe that the ℓth (previously basic) variable has now become zero, whereas the jth (previously nonbasic) variable has become positive. This suggests that $\boldsymbol{A}_j$ should replace $\boldsymbol{A}_{j_\ell}$ in the basis. That is, we replace the old basis matrix $\boldsymbol{B}$ with

$$\bar{\boldsymbol{B}} = [\, \boldsymbol{A}_{j_1}, \ldots, \boldsymbol{A}_{j_\ell - 1}, \boldsymbol{A}_j, \boldsymbol{A}_{j_\ell + 1}, \ldots, \boldsymbol{A}_{j_m} \,].$$

Equivalently, the basis indices $\{j_1, \ldots, j_m\}$ are replaced by $\{\bar{j}_1, \ldots, \bar{j}_m\}$, where

$$\bar{j}_i = \begin{cases} j_i & \text{if } i \neq \ell, \\ j & \text{if } i = \ell. \end{cases}$$

We say that x_{j_ℓ} (or $\boldsymbol{A}_{j_\ell}$) *exits the basis* and x_j (or $\boldsymbol{A}_j$) *enters the basis*.

- Let us revisit Problem 3.1 and consider the basic feasible solution $\boldsymbol{x} = (1, 1, 0, 0)$ for which we found the reduced cost of the nonbasic variable x_3 is $\bar{c}_3 = -\frac{3}{2}c_1 + \frac{1}{2}c_2 + c_3$. Suppose that $\boldsymbol{c} = (2, 0, 0, 0)$. Then $\bar{c}_3 = -3$ is negative, so we can reduce the cost by moving in the corresponding simplex direction $\boldsymbol{d} = (-3/2, 1/2, 1, 0)$. As t increases, the only component of $\boldsymbol{x} + t\boldsymbol{d}$ that decreases is the first one, since only $d_1 < 0$. The largest possible value of t is $t^* = -x_1/d_1 = 2/3$, which takes us to the point $\boldsymbol{y} = \boldsymbol{x} + 2\boldsymbol{d}/3 = (0, 4/3, 2/3, 0)$. At this point, the variable x_3 has entered the basis and the variable x_1 has exited. Note that the columns $\boldsymbol{A}_3 = (1, 3)$ (which replaces $\boldsymbol{A}_1$ in the basis matrix) and $\boldsymbol{A}_2 = (1, 0)$ are linearly independent and they therefore form a new basis matrix constructed from the column indices $\boldsymbol{j} = (3, 2)$:

$$\bar{\boldsymbol{B}} = \begin{bmatrix} 1 & 1 \\ 3 & 0 \end{bmatrix}.$$

  Notice also that $\boldsymbol{y}$ is the basic feasible solution corresponding to $\bar{\boldsymbol{B}}$:

$$\bar{\boldsymbol{B}}^{-1}\boldsymbol{b} = \frac{1}{-3}\begin{bmatrix} 0 & -1 \\ -3 & 1 \end{bmatrix}\begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 2/3 \\ 4/3 \end{bmatrix}.$$

  The following theorem establishes that these observations hold generally.

Theorem 3.2:

(i) *The columns* $\{\boldsymbol{A}_{\bar{j}_i} : i = 1, \ldots, m\}$ *are linearly independent, and hence* $\bar{\boldsymbol{B}}$ *is a basis matrix.*

(ii) *The vector* $\boldsymbol{y} = \boldsymbol{x} + t^*\boldsymbol{d}$ *is a basic feasible solution associated with the basis matrix* $\bar{\boldsymbol{B}}$.

Proof: (i) If the vectors $\{\boldsymbol{A}_{\bar{j}_i} : i = 1, \ldots, m\}$ were linearly dependent, then there would exist coefficients $\lambda_1, \ldots, \lambda_m$, not all zero, such that

$$\sum_{i=1}^{m} \lambda_i \boldsymbol{A}_{\bar{j}_i} = \boldsymbol{0}.$$

On multiplying this statement by $\boldsymbol{B}^{-1}$, the vectors $\boldsymbol{B}^{-1}\boldsymbol{A}_{\bar{j}_i}$ would then be seen to be linearly dependent. Since $\boldsymbol{B}^{-1}\boldsymbol{B} = \boldsymbol{1}$, where $\boldsymbol{1}$ is the $m \times m$ identity matrix, and the ith column of $\boldsymbol{B}$ is $\boldsymbol{A}_{j_i}$, we know that $\boldsymbol{B}^{-1}\boldsymbol{A}_{j_i}$ is just the ith unit vector $\boldsymbol{e}_i$, so that $\{\boldsymbol{B}^{-1}\boldsymbol{A}_{\bar{j}_i} : i \neq \ell\} = \{\boldsymbol{B}^{-1}\boldsymbol{A}_{j_i} : i \neq \ell\} = \{\boldsymbol{e}_i : i \neq \ell\}$ are linearly independent. Moreover, $\boldsymbol{B}^{-1}\boldsymbol{A}_j$ is linearly independent of these vectors: its ℓth component is the positive value $-d_{j_\ell}$ while their ℓth components are all zero. This is a contradiction.

(ii) We note that $\boldsymbol{y} \geq \boldsymbol{0}$, $\boldsymbol{A}\boldsymbol{y} = \boldsymbol{b}$, and $y_i = 0$ for $i \neq \bar{j}_1, \ldots, \bar{j}_m$. From (i), we know that the columns of $\bar{\boldsymbol{B}}$ are linearly independent. Thus, $\boldsymbol{y}$ is a basic feasible solution corresponding to $\bar{\boldsymbol{B}}$.

3.B. An iteration of the simplex method

1. Start with a basis consisting of the basic columns $\boldsymbol{A}_{j_1}, \ldots, \boldsymbol{A}_{j_m}$, and an associated basic feasible solution $\boldsymbol{x}$.

2. Compute the reduced costs $\bar{c}_j = c_j - \boldsymbol{c}_B^\intercal \boldsymbol{u}$ for all nonbasic indices j. If they are all nonnegative, the current basic feasible solution is optimal and the algorithm terminates. Otherwise choose some j for which $\bar{c}_j < 0$. Let $\boldsymbol{u} = -\boldsymbol{d}_B = \boldsymbol{B}^{-1}\boldsymbol{A}_j$, where $\boldsymbol{A}_j$ is the column that *enters the basis*.

3. If no component of $\boldsymbol{u}$ is positive, we have $t^* = \infty$, the optimal cost is $-\infty$ and the algorithm terminates.

4. If some component of $\boldsymbol{u}$ is positive, let
$$t^* = \min_{\substack{i=1,\ldots,m:\\ u_i>0}} \frac{x_{j_i}}{u_i}.$$

5. Let ℓ be such that $t^* = x_{j_\ell}/u_\ell$. Form a new basis by replacing $\boldsymbol{A}_{j_\ell}$ with $\boldsymbol{A}_j$. The new basic feasible solution $\boldsymbol{y}$ has components $y_j = t^*$ and $y_{j_i} = x_{j_i} - t^* u_i$ for $i \neq \ell$.

Theorem 3.3: *Assume that the feasible set is nonempty and that every basic feasible solution is nondegenerate. Then the simplex method terminates after a finite number of iterations. At termination there are two possibilities:*

1. *We have an optimal basis* $\mathbf{B}$ *and an associated optimal basic feasible solution.*

2. *We have found a vector* $\mathbf{d}$ *satisfying* $\mathbf{A}\mathbf{d} = \mathbf{0}, \mathbf{d} \geq \mathbf{0}$, *and* $\mathbf{c}^\intercal \mathbf{d} < 0$; *proceeding indefinitely in this direction leads to the optimal cost* $-\infty$.

3.C. The simplex method for degenerate problems

In a degenerate linear programming problem, the following possibilities may be encountered:

(a) If the current basic feasible solution $\boldsymbol{x}$ is degenerate, t^* can be zero, so that the new basic feasible solution $\boldsymbol{y}$ is the same as $\boldsymbol{x}$. This happens when $x_{j_\ell} = 0$ and $d_{j_\ell} < 0$. Nevertheless, we can still define the new basis $\bar{\boldsymbol{B}}$ by replacing $\boldsymbol{A}_{j_\ell}$ with $\boldsymbol{A}_j$, and Theorem 3.2 still holds.

(b) Even if t^* is positive, it may happen that more than one of the original basic variables becomes zero at the new point $\boldsymbol{x} + t^*\boldsymbol{d}$. Since only one of them exits the basis, the other zero values remain in the basis, so that the new basic feasible solution $\boldsymbol{y}$ is degenerate.

- In Figure 3.1 we visualize a feasible set in standard form, with $n - m = 2$, by standing on the two-dimensional plane defined by $\boldsymbol{Ax} = \boldsymbol{b}$.

Q. What kind of point is $\boldsymbol{x}$?

Degenerate basic feasible solution.

Q. Assume that x_4 and x_5 are the nonbasic variables. What are the corresponding simplex directions?

$\boldsymbol{g}$ and $\boldsymbol{f}$.

Q. What are the corresponding values of t^*?

0 and 0.

Q. If we perform a change of basis, with x_4 entering and x_6 exiting, what are the new nonbasic variables?

x_6 and x_5.

Q. What are their corresponding simplex directions?

$-\boldsymbol{g}$ and $\boldsymbol{h}$.

Q. If we follow the direction $\boldsymbol{h}$, can we reach a new basic feasible solution at a lower cost?

Yes, we will arrive at the new basic solution $\boldsymbol{y}$, which has a lower cost than $\boldsymbol{x}$.

A sequence of basis changes may lead back to the initial basis, in which case the algorithm may loop indefinitely. This undesirable phenomenon is called *cycling*.

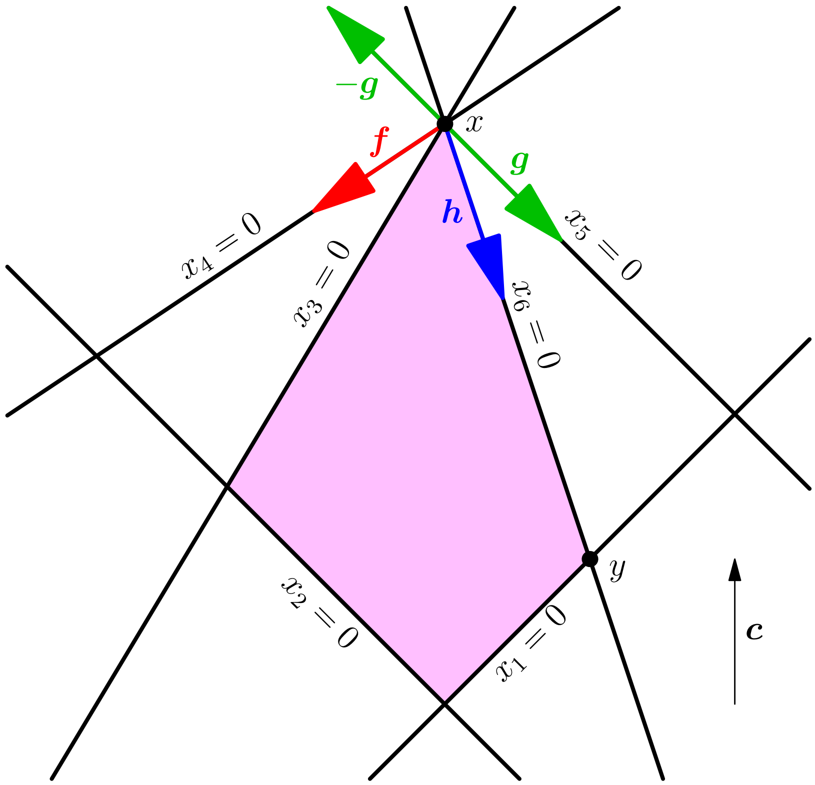


Figure 3.1:

3.D. Implementation of the simplex method

A number of implementations of the simplex method are available. Which implementation is most efficient for a given problem depends very much on the structure of the matrix $\boldsymbol{A}$, as well as the cost vector $\boldsymbol{c}$.

3.D.1 Naive implementation

If we carry no auxiliary information from one iteration to another, then there are three major computations to be handled at each iteration (a linear solve, up to $n-m$ dot products, followed by another linear solve):

$$\begin{aligned}
&\boldsymbol{p}^\intercal = \boldsymbol{c}_B^\intercal \boldsymbol{B}^{-1} \Rightarrow \boldsymbol{B}^\intercal \boldsymbol{p} = \boldsymbol{c}_B, \\
&\bar{c}_j = c_j - \boldsymbol{p}^\intercal \boldsymbol{A}_j \quad \text{for } j \in N, \\
&\boldsymbol{B}\boldsymbol{u} = \boldsymbol{A}_j,
\end{aligned}$$

where N is the set of nonbasic indices. However, depending on the implementation, one may not need to calculate $\bar{c}_j$ for every $j \in N$. If one wants to always follow the direction with the **most negative rate of cost change**, one must compute the reduced costs for every nonbasic variable. The same situation holds if one wants to follow the direction which leads to the **greatest cost reduction** $-t^*\bar{c}_j$. However, if one simply chooses the first variable encountered with a negative reduced cost, one need not compute the reduced costs for the remaining nonbasic variables. Because of this savings, one typically finds in practice that the latter choice is most efficient, even though one it doesn't necessarily follow the *path of steepest descent*. These implementation-dependent choices as to which variables enter and (at degenerate vertices) exit are known as *pivot rules*.

Once the entering variable x_j is known, we need to compute the displacement vector $\boldsymbol{u} = \boldsymbol{B}^{-1}\boldsymbol{A}_j$ that determines the direction of motion $\boldsymbol{d}$ and limiting parameter t^*. The computational cost of the two linear solves is $\mathcal{O}(m^3)$. If we assume that $n \gg m$, the computational cost of computing all nonbasic reduced costs is $\mathcal{O}(nm)$. The overall computational cost of a naive implementation is thus $\mathcal{O}(m^3 + nm)$.

3.D.2 Revised simplex method

Typically, the expensive steps of the naive implementation are the two linear solves. Since the matrix $\boldsymbol{B}$ appears in both linear systems, it may seem reasonable to first compute $\boldsymbol{B}^{-1}$ and then perform the two matrix–vector multiplies $\boldsymbol{c}_B^\intercal \boldsymbol{B}^{-1}$ and $\boldsymbol{B}^{-1}\boldsymbol{A}_j$. However, this alternative still requires $\mathcal{O}(m^3 + nm)$ operations for an entire iteration. Fortunately, there is a more efficient method for updating the matrix $\boldsymbol{B}^{-1}$ each time that we perform a change of basis, based on the previously calculated inverse. Recall that

$$\boldsymbol{B} = [\,\boldsymbol{A}_{j_1}, \ldots, \boldsymbol{A}_{j_m}\,]$$

and

$$\bar{\boldsymbol{B}} = [\,\boldsymbol{A}_{j_1}, \ldots, \boldsymbol{A}_{j_\ell - 1}, \boldsymbol{A}_j, \boldsymbol{A}_{j_\ell + 1}, \ldots, \boldsymbol{A}_{j_m}\,].$$

Note that

$$\boldsymbol{B}^{-1}\bar{\boldsymbol{B}} = [\,\boldsymbol{e}_1, \ldots, \boldsymbol{e}_{\ell-1}, \boldsymbol{u}, \boldsymbol{e}_{\ell+1}, \ldots, \boldsymbol{e}_m\,],$$

since $\boldsymbol{u} = \boldsymbol{B}^{-1}\boldsymbol{A}_j$.

Definition: The operation of adding a constant multiple of one row of a matrix to another row (or even the same row) is called an *elementary row operation*.

Remark: An elementary row operation on an $m \times n$ matrix $\boldsymbol{A}$ is equivalent to multiplication on the left by an $m \times m$ matrix $\boldsymbol{Q}$.

- Let

$$\boldsymbol{Q} = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad \boldsymbol{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$$

The matrix $\boldsymbol{QA}$ is the result of adding twice the second row of $\boldsymbol{A}$ to the first:

$$\boldsymbol{A} = \begin{bmatrix} 7 & 10 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}.$$

Let us now apply the following sequence of elementary row operations to $\boldsymbol{B}^{-1}\bar{\boldsymbol{B}}$:

1. Recalling that $u_\ell > 0$, add row ℓ times $-u_i/u_\ell$ to each row $i \neq \ell$, thereby reducing u_i to 0.
2. Divide row ℓ by u_ℓ. This sets u_ℓ to 1. Note that this is equivalent to adding row ℓ times $1/u_\ell - 1$ to itself.

The above operations are equivalent to multiplication on the left by a matrix $\boldsymbol{Q}$. We have chosen this sequence specifically to reduce $\boldsymbol{B}^{-1}\bar{\boldsymbol{B}}$ to the $m \times m$ identity matrix $\mathbf{1}$. That is,

$$\boldsymbol{Q}\boldsymbol{B}^{-1}\bar{\boldsymbol{B}} = \mathbf{1}.$$

Since

$$\boldsymbol{Q}\boldsymbol{B}^{-1} = \bar{\boldsymbol{B}}^{-1},$$

we now see that to compute $\bar{\boldsymbol{B}}^{-1}$, we only need to apply the above sequence of elementary row operations to $\boldsymbol{B}^{-1}$! This requires only $\mathcal{O}(m)$ arithmetic computations and leads to the following implementation, known as the *revised simplex method*:

1. In a typical iteration, we start with a basis consisting of the basic columns $\boldsymbol{A}_{j_1}, \ldots, \boldsymbol{A}_{j_m}$, an associated basic feasible solution $\boldsymbol{x}$, and the inverse $\boldsymbol{B}^{-1}$ of the basis matrix.
2. Compute the row vector $\boldsymbol{p}^\intercal = \boldsymbol{c}_B^\intercal \boldsymbol{B}^{-1}$ and the reduced costs $\bar{c}_j = c_j - \boldsymbol{p}^\intercal \boldsymbol{A}_j$. If they are all nonnegative, the current basic feasible solution is optimal, and the algorithm terminates; otherwise, choose some j for which $\bar{c}_j < 0$.
3. Compute $\boldsymbol{u} = \boldsymbol{B}^{-1}\boldsymbol{A}_j$. If no component of $\boldsymbol{u}$ is positive, the optimal cost is $-\infty$, and the algorithm terminates.
4. If some component of $\boldsymbol{u}$ is positive, let

$$t^* = \min_{\substack{i=1,\ldots,m: \\ u_i > 0}} \frac{x_{j_i}}{u_i}.$$

5. Let ℓ be such that $t^* = x_{j_\ell}/u_\ell$. Form a new basis by replacing $\boldsymbol{A}_{j_\ell}$ with $\boldsymbol{A}_j$. If $\boldsymbol{y}$ is the new basic feasible solution, the values of the new basic variables are $y_j = t^*$ and $y_{j_i} = x_{j_i} - t^* u_i$ for $i \neq \ell$.
6. Form the $m \times (m+1)$ matrix $[\boldsymbol{B}^{-1} | \boldsymbol{u}]$. Add to each of its rows a multiple of row ℓ to make the last column equal to the unit vector $\boldsymbol{e}_\ell$. The first m columns of the result is the matrix $\bar{\boldsymbol{B}}^{-1}$.

3.D.3 Full tableau implementation

Instead of maintaining and updating the matrix $\boldsymbol{B}^{-1}$, let us maintain and update the $m \times (n+1)$ matrix $\boldsymbol{B}^{-1}[\boldsymbol{b}|\boldsymbol{A}]$ with columns $\boldsymbol{B}^{-1}\boldsymbol{b}, \boldsymbol{B}^{-1}\boldsymbol{A}_1, \ldots, \boldsymbol{B}^{-1}\boldsymbol{A}_n$. One advantage of doing this is that one does not need to need to allocate a separate column for the vector $\boldsymbol{u} = \boldsymbol{B}^{-1}\boldsymbol{A}_j$, corresponding to the variable entering the basis. This special column is called the *pivot column*. If the ℓth basic variable exits the basis, the ℓth row is called the *pivot row*. The element corresponding to both pivot row and column is the *pivot element*. Note that the pivot element u_ℓ remains positive until the algorithm terminates.

We refer to the initial column $\boldsymbol{B}^{-1}\boldsymbol{b}$, which contains the components of the basic variable $\boldsymbol{x}_B$, as the *zeroth column*. One way to remember the form of this augmented matrix is to multiply the standard-form constraint $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x}$ by $\boldsymbol{B}^{-1}$ on each side:

$$\boldsymbol{B}^{-1}\boldsymbol{b} = \boldsymbol{B}^{-1}\boldsymbol{A}\boldsymbol{x};$$

the rows of the augmented matrix tabulate the coefficients of this equality constraint.

It is also convenient to add a *zeroth row* to represent the negative cost $-\boldsymbol{c}_B^\mathsf{T}\boldsymbol{x}_B$ in the zeroth column, followed by the reduced costs $\bar{c}_j$ of each of the n variables in subsequent columns, to obtain the *tableau*

$-\boldsymbol{c}_B^\mathsf{T}\boldsymbol{x}_B$	$\bar{c}_1$	$\ldots$	$\bar{c}_n$
$\boldsymbol{B}^{-1}\boldsymbol{b}$	$\boldsymbol{B}^{-1}\boldsymbol{A}_1$	$\ldots$	$\boldsymbol{B}^{-1}\boldsymbol{A}_n$

Remark: As we now show, the reason for the minus sign in the expression $-\boldsymbol{c}_B^\mathsf{T}\boldsymbol{x}_B$ in the zeroth column and zeroth row is for consistency with the reduced costs $\bar{c}_j = c_j - \boldsymbol{c}_B^\mathsf{T}\boldsymbol{B}^{-1}\boldsymbol{A}_j$. That is, the same update rule can then be used for all rows: a multiple of the pivot row is added to the zeroth row to set the reduced cost of the entering variable x_j to zero, as must be the case for a basic variable. At the beginning of the iteration, the zeroth row is of the form

$$[0|\boldsymbol{c}^\mathsf{T}] - \boldsymbol{p}^\mathsf{T}[\boldsymbol{b}|\boldsymbol{A}],$$

where $\boldsymbol{p}^\mathsf{T} = \boldsymbol{c}_B^\mathsf{T}\boldsymbol{B}^{-1}$. The term $\boldsymbol{p}^\mathsf{T}[\boldsymbol{b}|\boldsymbol{A}]$ is a linear combination of the rows of $[\boldsymbol{b}|\boldsymbol{A}]$. Since each row of $[\boldsymbol{B}^{-1}\boldsymbol{b}|\boldsymbol{B}^{-1}\boldsymbol{A}]$ is also a linear combination of the rows of $[\boldsymbol{b}|\boldsymbol{A}]$, adding a multiple of the pivot row to the zeroth row yields $[0|\boldsymbol{c}^\mathsf{T}]$ minus some linear combination $\bar{\boldsymbol{p}}^\mathsf{T}$ of the rows of $[\boldsymbol{b}|\boldsymbol{A}]$:

$$[0|\boldsymbol{c}^\mathsf{T}] - \bar{\boldsymbol{p}}^\mathsf{T}[\boldsymbol{b}|\boldsymbol{A}].$$

Our update rule is specifically chosen to annihilate the jth entry in the resulting zeroth row:

$$c_j - \bar{\boldsymbol{p}}^\mathsf{T}\boldsymbol{A}_j = 0.$$

Since $\boldsymbol{B}^{-1}\boldsymbol{A}_{j_i} = \boldsymbol{e}_i$, adding a multiple of the pivot row to the zeroth row will not change the reduced costs of the other basic variables x_{j_i} $(i \neq \ell)$ from their previous zero values. The reduced cost of each of the new basic variables is therefore zero:

$$\boldsymbol{c}_B^\mathsf{T} - \bar{\boldsymbol{p}}^\mathsf{T}\bar{\boldsymbol{B}} = \boldsymbol{0},$$

recalling that $\bar{\boldsymbol{B}} = [\boldsymbol{A}_{j_1}, \ldots, \boldsymbol{A}_{j_\ell - 1}, \boldsymbol{A}_j, \boldsymbol{A}_{j_\ell + 1}, \ldots, \boldsymbol{A}_{j_m}]$. We thus conclude that $\bar{\boldsymbol{p}}^\mathsf{T} = \boldsymbol{c}_B^\mathsf{T}\bar{\boldsymbol{B}}^{-1}$; this implies that the updated zeroth row equals

$$[0|\boldsymbol{c}^\mathsf{T}] - \boldsymbol{c}_B^\mathsf{T}\bar{\boldsymbol{B}}^{-1}[\boldsymbol{b}|\boldsymbol{A}],$$

as desired.

The full simplex tableau algorithm can now be summarized. Given a basis matrix $\boldsymbol{B}$, one initializes the full tableau with the basic components of the solution $\boldsymbol{x}_B = \boldsymbol{B}^{-1}\boldsymbol{b}$ and the matrix product $\boldsymbol{B}^{-1}\boldsymbol{A}$. The negative cost $-\boldsymbol{c}_B^\mathsf{T}\boldsymbol{x}_B$ and reduced cost vector $\bar{\boldsymbol{c}}^\mathsf{T} = \boldsymbol{c}^\mathsf{T} - \boldsymbol{c}_B^\mathsf{T}\boldsymbol{B}^{-1}\boldsymbol{A}$ are entered in the zeroth row of the tableau. A single iteration of the *full simplex tableau* involves these four steps:

1. If the reduced costs are all nonnegative, the current basic feasible solution is optimal, and the algorithm terminates.

2. Otherwise, choose some j for which $\bar{c}_j < 0$. Denote column j (the pivot column) of the tableau by $\boldsymbol{u}$. If no component of $\boldsymbol{u}$ is positive, the optimal cost is $-\infty$, and the algorithm terminates.

3. For each i for which u_i is positive, compute the ratio x_{j_i}/u_i. Let ℓ be the index of a row that corresponds to the smallest ratio. The column $\boldsymbol{A}_{j_\ell}$ exits the basis and the column $\boldsymbol{A}_j$ enters the basis.

4. Add a constant multiple of row ℓ (the pivot row) so that u_ℓ (the pivot element) becomes 1 and all other entries of the pivot column become 0.

Problem 3.3: Use the simplex method to solve the linear programming problem

$$\begin{array}{lrcrcrcl}
\text{minimize} & -10x_1 & - & 12x_2 & - & 12x_3 & & \\
\text{subject to} & x_1 & + & 2x_2 & + & 2x_3 & \leq & 20, \\
& 2x_1 & + & x_2 & + & 2x_3 & \leq & 20, \\
& 2x_1 & + & 2x_2 & + & x_3 & \leq & 20, \\
& \multicolumn{5}{l}{x_1, x_2, x_3 \geq 0.}
\end{array}$$

First, we introduce slack variables x_4, x_5, x_6 to form the standard-form problem

$$\begin{array}{lrcrcrcrcrcrcl}
\text{minimize} & -10x_1 & - & 12x_2 & - & 12x_3 & & & & & & & & \\
\text{subject to} & x_1 & + & 2x_2 & + & 2x_3 & + & x_4 & & & & & = & 20, \\
& 2x_1 & + & x_2 & + & 2x_3 & & & + & x_5 & & & = & 20, \\
& 2x_1 & + & 2x_2 & + & x_3 & & & & & + & x_6 & = & 20, \\
& \multicolumn{13}{l}{x_1, x_2, x_3, x_4, x_5, x_6 \geq 0.}
\end{array}$$

We need a basic feasible solution to initialize the simplex method. An obvious basic solution is $(0, 0, 0, 20, 20, 20)$, with $j_1 = 4$, $j_2 = 5$, $j_3 = 6$, and the identity basis matrix. This corresponds to the vertex $A = (0, 0, 0)$ of the original problem depicted in Figure 3.2. Since $\boldsymbol{c}_B = \boldsymbol{0}$, the zeroth row of the initial tableau contains the negative cost $-\boldsymbol{c}_B^\mathsf{T}\boldsymbol{x}_B = 0$ followed by the elements of the reduced costs $\bar{\boldsymbol{c}} = \boldsymbol{c}$:

			x_1^*	x_2	x_3	x_4	x_5	x_6
		0	-10	-12	-12	0	0	0
x_4	=	20	1	2	2	1	0	0
$x_5^\dagger$	=	20	2	1	2	0	1	0
x_6	=	20	2	2	1	0	0	1

The decision variables corresponding to each column are tabulated above the zeroth row. The basic variables are listed next to each of their values in the zeroth column. We are now ready to begin the first simplex iteration.

Let us use the first column with a negative reduced cost, x_1 as our entering variable. We highlight the corresponding pivot column with an asterisk and let $\boldsymbol{u} = (1, 2, 2)$. We then compute the ratios x_{j_i}/u_i for each i such that $u_i > 0$. These ratios are $20/1 = 20$, $20/2 = 10$, $20/2 = 10$, respectively. For the exit variable, we choose the first variable, x_5, that achieves the minimum ratio (10), highlighting the pivot row with a dagger. We then apply elementary operations to the matrix to reduce the pivot column to $(0, 0, 1, 0)$:

			x_1	x_2^*	x_3	x_4	x_5	x_6
		100	0	-7	-2	0	5	0
x_4	=	10	0	$3/2$	1	1	$-1/2$	0
x_1	=	10	1	$1/2$	1	0	$1/2$	0
$x_6^\dagger$	=	0	0	1	-1	0	-1	1

This corresponds to the degenerate vertex $\boldsymbol{D} = (10, 0, 0)$ of the original problem depicted in Figure 3.2. The basic variables are now x_4, x_1, and x_6. In the tableau, the degeneracy is apparent from the value $x_6 = 0$. We now repeat the iteration, choosing x_2 as our entering variable and letting $\boldsymbol{u} = (3/2, 1/2, 1)$. We compute $x_4/u_1 = 10/(3/2) = 20/3$, $x_1/u_2 = 10/(1/2) = 20$, $x_6/u_3 = 0/1 = 0$. We therefore choose x_6 as our exiting variable and row reduce to obtain $(0, 0, 0, 1)$ in the pivot column. This row reduction accomplishes a change of basis but leaves us still at the point $\boldsymbol{D} = (10, 0, 0)$:

			x_1	x_2	x_3^*	x_4	x_5	x_6
		100	0	0	-9	0	-2	7
$x_4^\dagger$	=	10	0	0	$5/2$	1	1	$-3/2$
x_1	=	10	1	0	$3/2$	0	1	$-1/2$
x_2	=	0	0	1	-1	0	-1	1

Our basic variables are now x_4, x_1, and x_2. Next, we choose x_3 as our entering variable and let $\boldsymbol{u} = (5/2, 3/2, -1)$. We compute $x_4/u_1 = 10/(5/2) = 4$ and $x_1/u_2 = 10/(3/2) = 20/3$. The exiting variable is thus x_4. Upon row reducing, we obtain a zeroth row containing no negative reduced costs:

			x_1	x_2	x_3	x_4	x_5	x_6
		136	0	0	0	18/5	8/5	8/5
x_3	$=$	4	0	0	1	2/5	2/5	$-3/5$
x_1	$=$	4	1	0	0	$-3/5$	2/5	2/5
x_2	$=$	4	0	1	0	2/5	$-3/5$	2/5

We have thus reached the optimal solution $E = (4, 4, 4)$, for which the optimal cost is -136.

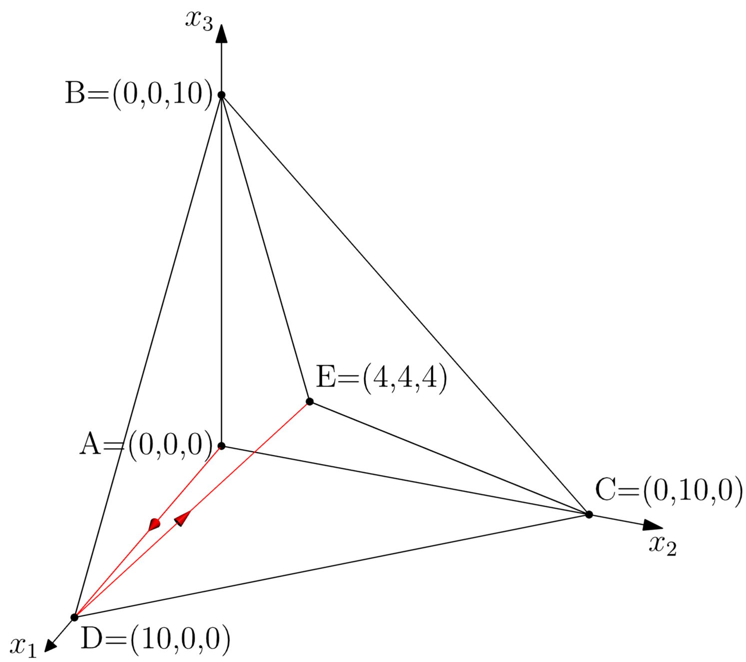


Figure 3.2: Simplex iterations.

Problem 3.4: Use the simplex method to solve the standard-form linear programming problem

$$\begin{array}{lrcrcrcrcl} \text{minimize} & x_1 & + & 3x_2 & + & 2x_3 & & & & \\ \text{subject to} & 2x_1 & - & x_2 & + & x_3 & & & = & 6, \\ & x_1 & - & x_2 & & & - & x_4 & = & 2, \\ & \multicolumn{9}{l}{x_1, x_2, x_3, x_4 \geq 0.} \end{array}$$

We need a basic feasible solution to initialize the tableau. The choice $j_1 = 1$ and $j_2 = 2$ corresponds to

$$\boldsymbol{B} = \begin{bmatrix} 2 & -1 \\ 1 & -1 \end{bmatrix}, \qquad \boldsymbol{x}_B = \frac{1}{-1}\begin{bmatrix} -1 & 1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 6 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \end{bmatrix},$$

and

$$\boldsymbol{B}^{-1}\boldsymbol{A} = \frac{1}{-1}\begin{bmatrix} -1 & 1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 2 & -1 & 1 & 0 \\ 1 & -1 & 0 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 2 \end{bmatrix}.$$

Let $\boldsymbol{c} = (1,3,2,0)$, so that $\boldsymbol{c}_B = (1,3)$. We can use the last two columns of the above result to easily calculate the reduced costs of the nonbasic variables:

$$\bar{c}_3 = c_3 - \boldsymbol{c}_B^{\mathsf{T}}\boldsymbol{B}^{-1}\boldsymbol{A}_3 = 2 - [1 \quad 3]\begin{bmatrix} 1 \\ 1 \end{bmatrix} = -2,$$

$$\bar{c}_4 = c_4 - \boldsymbol{c}_B^{\mathsf{T}}\boldsymbol{B}^{-1}\boldsymbol{A}_4 = 0 - [1 \quad 3]\begin{bmatrix} 1 \\ 2 \end{bmatrix} = -7.$$

We enter these results, along with the negative cost at the initial basic feasible solution, $-\boldsymbol{c}_B^{\mathsf{T}}\boldsymbol{x}_B = -10$, in the tableau:

			x_1	x_2	x_3^*	x_4
		-10	0	0	-2	-7
x_1	$=$	4	1	0	1	1
$x_2^\dagger$	$=$	2	0	1	1	2

Two iterations of the simplex method then bring us to the optimal solution $(3,0,0,1)$ with optimal cost 3:

			x_1	x_2	x_3	x_4^*
		-6	0	2	0	-3
x_1	$=$	2	1	-1	0	-1
$x_3^\dagger$	$=$	2	0	1	1	2

			x_1	x_2	x_3	x_4
		-3	0	$7/2$	$3/2$	0
x_1	$=$	3	1	$-1/2$	$1/2$	0
x_4	$=$	1	0	$1/2$	$1/2$	1

Remark: The entering decision variable always has a negative reduced cost. Often there is more than one choice for this entering variable. In the previous examples of the full simplex tableau, we always chose the entering variable to be the one with the lowest subscript. This is particularly advantageous in implementations of the revised simplex method, where the reduced costs are computed separately from the tableau. If one computes the reduced costs sequentially, starting with the lowest subscript, one can stop computing reduced costs as soon as the first negative value is encountered. Although the full simplex tableau is perhaps more elegant, in that the reduced cost computation is incorporated directly into the row operations, calculating the reduced costs in this iterative manner has the disadvantage that all reduced costs must be calculated (no matter what pivot rule is used), so that the reduced costs of the other variables are available for future iterations.

Remark: For both the revised and full tableau simplex methods, there is nevertheless an important advantage in always choosing the entering variable with the **lowest subscript** that has a negative reduced cost. If we also choose the exiting variable to be the variable with the **lowest subscript** j_i that achieves the minimum ratio x_{j_i}/u_i, the simplex method never cycles and will always terminate after a finite number of iterations. This widely used pivot rule is known as *Bland's rule*.

Problem 3.5: Consider the linear programming problem

$$\begin{array}{lrcrcrcrcrcrcrcl}
\text{minimize} & -3x_1 & + & 80x_2 & - & 14x_3 & + & 24x_4 & & & & & - & 12x_7 & & \\
\text{subject to} & \frac{1}{4}x_1 & - & 8x_2 & - & x_3 & + & 9x_4 & + & x_5 & & & & & = & 0, \\
 & \frac{1}{2}x_1 & - & 12x_2 & - & \frac{1}{2}x_3 & + & 3x_4 & & & + & x_6 & & & = & 0, \\
 & & & & & x_3 & & & & & & & + & x_7 & = & 4,
\end{array}$$

$$x_1, x_2, x_3, x_4, x_5, x_6, x_7 \geq 0.$$

An obvious basic solution is $(0,0,0,0,0,0,4)$, with $j_1 = 5$, $j_2 = 6$, $j_3 = 7$, and the identity basis matrix. Since $\boldsymbol{c}_B = (0,0,-12)$, the zeroth row of the initial tableau contains the negative cost $-\boldsymbol{c}_B^\intercal \boldsymbol{x}_B = 48$ followed by the elements of the reduced cost vector $[-3, 80, -14, 24, 0, 0, -12] - [0, 0, -12]\boldsymbol{A}$:

		x_1^*	x_2	x_3	x_4	x_5	x_6	x_7
	48	-3	80	-2	24	0	0	0
$x_5^\dagger =$	0	1/4	-8	-1	9	1	0	0
$x_6 =$	0	1/2	-12	$-1/2$	3	0	1	0
$x_7 =$	4	0	0	1	0	0	0	1

If we pivot according to Bland's rule, the simplex method reaches an optimal solution in six iterations:

		x_1	x_2^*	x_3	x_4	x_5	x_6	x_7
	48	0	-16	-14	132	12	0	0
$x_1 =$	0	1	-32	-4	36	4	0	0
$x_6^\dagger =$	0	0	4	3/2	-15	-2	1	0
$x_7 =$	4	0	0	1	0	0	0	1

		x_1	x_2	x_3^*	x_4	x_5	x_6	x_7
	48	0	0	-8	72	4	4	0
$x_1^\dagger =$	0	1	0	8	-84	-12	8	0
$x_2 =$	0	0	1	3/8	$-15/4$	$-1/2$	1/4	0
$x_7 =$	4	0	0	1	0	0	0	1

			x_1	x_2	x_3	x_4^*	x_5	x_6	x_7
		48	1	0	0	−12	−8	12	0
x_3	=	0	1/8	0	1	−21/2	−3/2	1	0
$x_2^\dagger$	=	0	−3/64	1	0	3/16	1/16	−1/8	0
x_7	=	4	−1/8	0	0	21/2	3/2	−1	1

			x_1^*	x_2	x_3	x_4	x_5	x_6	x_7
		48	−2	64	0	0	−4	4	0
x_3	=	0	−5/2	56	1	0	2	−6	0
x_4	=	0	−1/4	16/3	0	1	1/3	−2/3	0
$x_7^\dagger$	=	4	5/2	−56	0	0	−2	6	1

			x_1	x_2	x_3	x_4	x_5^*	x_6	x_7
		256/5	0	96/5	0	0	−28/5	44/5	4/5
x_3	=	4	0	0	1	0	0	0	1
$x_4^\dagger$	=	2/5	0	−4/15	0	1	2/15	−1/15	1/10
x_1	=	8/5	1	−112/5	0	0	−4/5	12/5	2/5

			x_1	x_2	x_3	x_4	x_5	x_6	x_7
		68	0	8	0	42	0	6	5
x_3	=	4	0	0	1	0	0	0	1
x_5	=	3	0	−2	0	15/2	1	−1/2	3/4
x_1	=	4	1	−24	0	6	0	2	1

However, suppose we always select the entering variable with the most negative reduced cost. The first four iterations will be exactly as above, until we arrive at

			x_1	x_2	x_3	x_4	x_5^*	x_6	x_7
		48	−2	64	0	0	−4	4	0
$x_3^\dagger$	=	0	−5/2	56	1	0	2	−6	0
x_4	=	0	−1/4	16/3	0	1	1/3	−2/3	0
x_7	=	4	5/2	−56	0	0	−2	6	1

After two further iterations using the most-negative reduced cost, we arrive back at the initial tableau!

			x_1	x_2	x_3	x_4	x_5	x_6^*	x_7
		48	−7	176	2	0	0	−8	0
x_5	=	0	−5/4	28	1/2	0	1	−3	0
$x_4^\dagger$	=	0	1/6	−4	−1/6	1	0	1/3	0
x_7	=	4	0	0	1	0	0	0	1

			x_1^*	x_2	x_3	x_4	x_5	x_6	x_7
		48	−3	80	−2	24	0	0	0
$x_5^\dagger$	=	0	1/4	−8	−1	9	1	0	0
x_6	=	0	1/2	−12	−1/2	3	0	1	0
x_7	=	4	0	0	1	0	0	0	1

Notice that at each iteration, we always remain at the degenerate vertex $(0, 0, 0, 0, 0, 0, 4)$, with nonoptimal cost -48. Although a change of basis is performed at each iteration, eventually we return to the initial basis. With this pivot rule, the simplex method cycles forever, never terminating. This example of *cycling* emphasizes the importance of adopting Bland's rule in practical implementations of the simplex method.

3.E. Finding an initial basic feasible solution

We now describe an efficient procedure for finding a basic feasible solution and associated basis matrix to initialize the simplex tableau.

First, multiply by -1 any equality constraints that have a negative right-hand side, to ensure that the standard-form problem

$$\begin{aligned} \text{minimize} \quad & \boldsymbol{c}^\mathsf{T}\boldsymbol{x} \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \\ & \boldsymbol{x} \geq \boldsymbol{0} \end{aligned}$$

satisfies $\boldsymbol{b} \geq \boldsymbol{0}$.

Next, introduce a vector $\boldsymbol{y} \in \mathbb{R}^m$ of *artificial variables* and consider the *auxiliary problem*

$$\begin{aligned} \text{minimize} \quad & \sum_{i=1}^{m} y_i \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{b}, \\ & \boldsymbol{x}, \boldsymbol{y} \geq \boldsymbol{0}. \end{aligned}$$

Remark: If $\boldsymbol{x}^*$ is a feasible solution to the given problem, $(\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{x}^*, \boldsymbol{0})$ yields an optimal zero-cost solution to the auxiliary problem. Hence, if the optimal cost of the auxiliary problem is positive, the given problem is *infeasible*.

Remark: On the other hand, a zero-cost basic feasible solution $(\boldsymbol{x}, \boldsymbol{y})$ to the auxiliary problem satisfies $\boldsymbol{y} = \boldsymbol{0}$, so that $\boldsymbol{x}$ is then a feasible solution to the given problem. If the associated basis matrix $\boldsymbol{B}$ contains only columns of $\boldsymbol{A}$, the columns in the tableau corresponding to the artificial variables $\boldsymbol{y}$ can simply be dropped.

Remark: A basic feasible solution to the auxiliary problem is given by $(\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{0}, \boldsymbol{b})$, with an identity basis matrix. Since all elements of $\boldsymbol{c}_B$ are one, the reduced cost of a nonbasic variable $\boldsymbol{x}_j$, which doesn't contribute to the auxiliary cost, is just the negative of the column sum of $\boldsymbol{A}_j$: $-\sum_{i=1}^{m} A_{i,j}$.

Remark: In Prob. 3.4, we solved the linear programming problem

$$\begin{aligned} \text{minimize} \quad & x_1 + x_2 + x_3 \\ \text{subject to} \quad & 2x_1 - x_2 + x_3 = 6, \\ & x_1 - x_2 - x_4 = 2, \\ & x_1, \; x_2, \; x_3, \; x_4 \ge 0 \end{aligned}$$

starting from a given basis matrix and corresponding basic feasible solution. Note that the right-hand side vector $\boldsymbol{b} = (6, 2)$ is already nonnegative. A systematic way to find an initial basis matrix and tableau is to introduce artificial variables y_1 and y_2 and consider the auxiliary problem

$$\begin{aligned} \text{minimize} \quad & y_1 + y_2 \\ \text{subject to} \quad & 2x_1 - x_2 + x_3 + y_1 = 6, \\ & x_1 - x_2 - x_4 + y_2 = 2, \\ & x_1, \; x_2, \; x_3, \; x_4, \; y_1, \; y_2 \ge 0. \end{aligned}$$

We start iterating from the initial solution $(0, 0, 0, 0, 6, 2)$ and identity basis matrix:

	x_1^*	x_2	x_3	x_4	y_1	y_2
-8	-3	2	-1	1	0	0
$y_1 =$ 6	2	-1	1	0	1	0
$y_2^\dagger =$ 2	1	-1	0	-1	0	1

Two iterations of the simplex method then lead to an optimal zero-cost solution of the auxiliary problem:

	x_1	x_2^*	x_3	x_4	y_1	y_2
-2	0	-1	-1	-2	0	3
$y_1^\dagger =$ 2	0	1	1	2	1	-2
$x_1 =$ 2	1	-1	0	-1	0	1

	x_1	x_2	x_3	x_4
0	0	0	0	0
$x_2 =$ 2	0	1	1	2
$x_1 =$ 4	1	0	1	1

In the final tableau, since none of the final basic variables are artificial variables, we have dropped their associated columns. Except for the zeroth row and the order of the basic variables (the order of the rows in the tableau is irrelevant), the resulting truncated tableau is identical to the initial tableau in Prob. 3.4. The tabulated columns of $\boldsymbol{B}^{-1}\boldsymbol{A}$ can then be used to calculate the reduced costs in the zeroth row of the initial tableau of the given problem.

Remark: In applying Bland's rule, any x variable should be considered to have a **lower subscript** than any y variable (e.g. in a tie, x_2 appears lexically before y_1).

Remark: One can sometimes reduce the number of artificial columns by exploiting the structure of $\boldsymbol{A}$: if a column $\boldsymbol{A}_j$ is a **positive** multiple of a unit vector $\boldsymbol{e}_k$, one can omit the column associated with the artificial variable y_k.

Remark: In a standard-form problem with m slack (not surplus) variables and $\boldsymbol{b} \geq \boldsymbol{0}$, the artificial columns can thus be eliminated completely, leaving one with an obvious basic feasible solution, as we saw in Prob. 3.3.

Problem 3.6: Let us optimize our search for an initial basic feasible solution for Prob. 3.4 by exploiting the fact that $\boldsymbol{A}_3 = \boldsymbol{e}_1$; this allows us to drop the column associated with y_1:

$$
\begin{array}{lrrrrrrl}
\text{minimize} & & & & & y_1 + & y_2 & \\
\text{subject to} & 2x_1 - & x_2 + & x_3 & & + y_1 & & = 6, \\
 & x_1 - & x_2 & & - x_4 & & + y_2 & = 2, \\
 & x_1, & x_2, & x_3, & x_4, & y_1, & y_2 & \geq 0.
\end{array}
$$

		x_1^*	x_2	x_3	x_4	y_2
	-8	-3	2	-1	1	0
$y_1=$	6	2	-1	1	0	0
$y_2^\dagger=$	2	1	-1	0	-1	1

		x_1	x_2^*	x_3	x_4	y_2
	-2	0	-1	-1	-2	3
$y_1^\dagger=$	2	0	1	1	2	-2
$x_1=$	2	1	-1	0	-1	1

		x_1	x_2	x_3	x_4
	0	0	0	0	0
$x_2=$	2	0	1	1	2
$x_1=$	4	1	0	1	1

Since the remaining artificial variable y_2 does not appear in the basis, we have dropped it from this zero-cost solution to the auxilliary problem.

Remark: The situation is more complicated if some of the artificial variables $\boldsymbol{y}$ are in the final basis $\boldsymbol{B}$. In this case, we must drive the artificial variables out of the basis. Let $I = \{i : x_{j_i} \text{ is a basic variable}\}$ denote the set of row indices in the zero-cost tableau that correspond to nonartificial basic variables. Suppose that the basic variable in the ℓth row of the tableau is an artificial variable, where $\ell \notin I$. If the ℓth row of $\boldsymbol{B}^{-1}\boldsymbol{A}$ contains only zeroes, then the matrix $\boldsymbol{A}$ has linearly dependent rows; we can therefore remove that row and continue. Otherwise, for some j the ℓth entry of $\boldsymbol{B}^{-1}\boldsymbol{A}_j$ is nonzero. Since $\boldsymbol{B}^{-1}\boldsymbol{A}_{j_i} = \boldsymbol{e}_i$ for $i \in I$, and $\ell \notin I$, we know that the ℓth entry of these vectors is zero. This means that $\boldsymbol{B}^{-1}\boldsymbol{A}_j$ is not a linear combination of these vectors, so that $\boldsymbol{A}_j$ is linearly independent of $\{\boldsymbol{A}_{j_i} : i \in I\}$. We can therefore allow x_j to enter the basis, with the artificial variable exiting. We can accomplish this by our usual row reduction. Although the pivot element (the ℓth entry of the jth column) could be negative, this doesn't affect the mechanics of the row operations; it is only important that the pivot element be nonzero.

Remark: Once we have arrived at a zero-cost optimal solution, we drive each artificial variable out of the basis, in the order they appear in the tableau (top to bottom), by iterating on the first **nonzero** pivot element of the corresponding row of $\boldsymbol{B}^{-1}\boldsymbol{A}$. If there is no nonzero pivot element (do not consider artificial columns), we delete that entire row from the tableau.

Problem 3.7: Find a basic feasible solution for the linear programming problem

$$\begin{array}{lrrrrrl}
\text{minimize} & -5x_1 - & x_2 + & 12x_3 + & x_4 & & \\
\text{subject to} & 3x_1 + & 2x_2 + & 2x_3 & & & = 10, \\
& 5x_1 + & 3x_2 + & x_3 + & x_4 & & = 16, \\
& x_1 + & x_2 + & x_3 & & -\ 2x_5 & = \ 5, \\
& x_1, & x_2, & x_3, & x_4, & x_5 & \geq \ 0.
\end{array}$$

Step 1. Multiply every constraint with a negative right-hand side by -1, so that $\boldsymbol{b} \geq \boldsymbol{0}$. In this case, the constraints are already in this form.

Step 2. In order to find a basic feasible solution, the artificial variables y_1, y_2, and y_3 are introduced:

$$\begin{array}{lrrrrrrrrl}
\text{minimize} & y_1 + & y_2 + & y_3 & & & & & & \\
\text{subject to} & 3x_1 + & 2x_2 + & 2x_3 & & & +\ y_1 & & & = 10, \\
& 5x_1 + & 3x_2 + & x_3 + & x_4 & & & +\ y_2 & & = 16, \\
& x_1 + & x_2 + & x_3 & & -\ 2x_5 & & & +\ y_3 & = \ 5, \\
& x_1, & x_2, & x_3, & x_4, & x_5, & y_1, & y_2, & y_3 & \geq \ 0.
\end{array}$$

Since the columns in $\boldsymbol{A}$ corresponding to x_4 and y_2 are identical, we can omit the column associated with the artificial variable y_2:

		x_1^*	x_2	x_3	x_4	x_5	y_1	y_3
	-31	-9	-6	-4	-1	2	0	0
$y_1=$	10	3	2	2	0	0	1	0
$y_2^\dagger=$	16	5	3	1	1	0	0	0
$y_3=$	5	1	1	1	0	-2	0	1

		x_1	x_2^*	x_3	x_4	x_5	y_1	y_3
	$-11/5$	0	$-3/5$	$-11/5$	$4/5$	2	0	0
$y_1^\dagger=$	$2/5$	0	$1/5$	$7/5$	$-3/5$	0	1	0
$x_1=$	$16/5$	1	$3/5$	$1/5$	$1/5$	0	0	0
$y_3=$	$9/5$	0	$2/5$	$4/5$	$-1/5$	-2	0	1

		x_1	x_2	x_3	x_4^*	x_5	y_1	y_3
	-1	0	0	2	-1	2	3	0
$x_2=$	2	0	1	7	-3	0	5	0
$x_1^\dagger=$	2	1	0	-4	2	0	-3	0
$y_3=$	1	0	0	-2	1	-2	-2	1

		x_1^*	x_2	x_3	x_4	x_5
	0	$1/2$	0	0	0	2
$x_2=$	5	$3/2$	1	1	0	0
$x_4=$	1	$1/2$	0	-2	1	0
$y_3^\dagger=$	0	$-1/2$	0	0	0	-2

After 3 iterations, we found an optimal (zero-cost) solution of the auxiliary problem, so we dropped the columns associated with the artificial variables. However, we still need to drive the variable y_3 out of the final basis, with x_1 entering:

		x_1	x_2	x_3	x_4	x_5
	0	0	0	0	0	0
$x_2=$	5	0	1	1	0	-6
$x_4=$	1	0	0	-2	1	-2
$x_1=$	0	1	0	0	0	4

We have thus obtained a basic feasible solution $(0,5,0,1,0)$ and corresponding values of $\boldsymbol{B}^{-1}\boldsymbol{A}$.

Problem 3.8: Find a basic feasible solution for the linear programming problem

$$\begin{array}{ll} \text{minimize} & 2x_1 + 6x_2 + x_3 + x_4 \\ \text{subject to} & x_1 + 2x_2 + x_4 = 6, \\ & x_1 + 2x_2 + x_3 + x_4 = 7, \\ & x_1 + 3x_2 - x_3 + 2x_4 = 7, \\ & x_1 + x_2 + x_3 = 5, \\ & x_1, x_2, x_3, x_4 \geq 0. \end{array}$$

Step 1. Multiply every constraint with a negative right-hand side by -1, so that $\boldsymbol{b} \geq \boldsymbol{0}$. In this case, there is no need to change the constraints.

Step 2. In order to find a basic feasible solution, the artificial variables y_1, y_2, y_3, and y_4 are introduced to form the auxiliary problem

$$\begin{array}{ll} \text{minimize} & y_1 + y_2 + y_3 + y_4 \\ \text{subject to} & x_1 + 2x_2 + x_4 + y_1 = 6, \\ & x_1 + 2x_2 + x_3 + x_4 + y_2 = 7, \\ & x_1 + 3x_2 - x_3 + 2x_4 + y_3 = 7, \\ & x_1 + x_2 + x_3 + y_4 = 5, \\ & x_1, x_2, x_3, x_4, y_1, y_2, y_3, y_4 \geq 0. \end{array}$$

		x_1^*	x_2	x_3	x_4	y_1	y_2	y_3	y_4
	-25	-4	-8	-1	-4	0	0	0	0
$y_1=$	6	1	2	0	1	1	0	0	0
$y_2=$	7	1	2	1	1	0	1	0	0
$y_3=$	7	1	3	-1	2	0	0	1	0
$y_4^\dagger=$	5	1	1	1	0	0	0	0	1

		x_1	x_2^*	x_3	x_4	y_1	y_2	y_3	y_4
	-5	0	-4	3	-4	0	0	0	4
$y_1^\dagger=$	1	0	1	-1	1	1	0	0	-1
$y_2=$	2	0	1	0	1	0	1	0	-1
$y_3=$	2	0	2	-2	2	0	0	1	-1
$x_1=$	5	1	1	1	0	0	0	0	1

		x_1	x_2	x_3^*	x_4	y_1	y_2	y_3	y_4
	-1	0	0	-1	0	4	0	0	0
$x_2=$	1	0	1	-1	1	1	0	0	-1
$y_2^\dagger=$	1	0	0	1	0	-1	1	0	0
$y_3=$	0	0	0	0	0	-2	0	1	1
$x_1=$	4	1	0	2	-1	-1	0	0	2

		x_1	x_2	x_3	x_4
	0	0	0	0	0
$x_2=$	2	0	1	0	1
$x_3=$	1	0	0	1	0
$y_3=$	0	0	0	0	0
$x_1=$	2	1	0	0	-1

We have arrived at an optimal solution of the auxiliary problem. However, we need to drive out the third variable in the basis, y_3, since it is an artificial variable. Since the third elements of $\boldsymbol{B}^{-1}\boldsymbol{A}_1, \ldots \boldsymbol{B}^{-1}\boldsymbol{A}_4$ are all zero, a basic feasible solution $(2, 2, 1, 0)$ to the original problem is obtained by simply dropping row 3:

		x_1	x_2	x_3	x_4
	0	0	0	0	0
$x_2=$	2	0	1	0	1
$x_3=$	1	0	0	1	0
$x_1=$	2	1	0	0	-1

3.F. The two-phase simplex method

The above considerations lead us to the so-called *two-phase simplex method*. In the first phase, the auxiliary problem is solved to find an initial tableau for the second phase. The second phase is the application of the simplex method on the given problem, starting from the basic feasible solution found in the first phase.

Phase I:

1. Multiply the constraints as needed by -1 so that $\boldsymbol{b} \geq \mathbf{0}$.

2. If a basic feasible solution is not known, introduce nonnegative artificial variables $y_1, \ldots, y_m$ and apply the simplex method to the auxiliary problem with cost $\sum_{i=1}^{m} y_i$.

3. If the optimal cost in the auxiliary problem is positive, the given problem is infeasible and the algorithm terminates.

4. If the optimal cost in the auxiliary problem is zero, a feasible solution to the given problem has been found. If no artificial variable is in the final basis, the artificial variables and the corresponding columns are eliminated, and a feasible basis for the given problem is available.

5. If the ℓth basic variable is an artificial one, examine the ℓth entry of the columns $\boldsymbol{B}^{-1}\boldsymbol{A}_j$, $j = 1, \ldots, n$. If all of these entries are zero, remove the ℓth row, as it represents a redundant constraint. Otherwise, if the ℓth entry of the jth column is nonzero, apply a change of basis (with this entry serving as the pivot element): the ℓth basic variable exits and x_j enters the basis. Repeat this procedure until all artificial variables are driven out of the basis.

Phase II:

1. Let the final basis and tableau obtained from Phase I be the initial basis and tableau for Phase II.

2. Compute the reduced costs of all variables for this initial basis, using the cost coefficients of the given problem and the tabulated values of $\boldsymbol{B}^{-1}\boldsymbol{A}$.

3. Apply the simplex method to the given problem.

Problem 3.9: Use the two-phase simplex method to determine the initial tableau for the linear programming problem

$$\begin{array}{rrrrl} \text{minimize} & x_1 & & -\,x_3 & \\ \text{subject to} & x_1\,+ & x_2 & & = \;\;4, \\ & & -x_2\,+ & x_3 & = -1, \\ & x_1, & x_2, & x_3 & \geq \;\;0. \end{array}$$

Phase I:
Step 1. Multiply every constraint with a negative right-hand side by -1, so that $\boldsymbol{b} \geq \boldsymbol{0}$:

$$\begin{array}{lrcrcrl} \text{minimize} & x_1 & & & - & x_3 & \\ \text{subject to} & x_1 & + & x_2 & & & = 4, \\ & & & x_2 & - & x_3 & = 1, \\ & x_1, & & x_2, & & x_3 & \geq 0. \end{array}$$

Step 2. In order to find a basic feasible solution, the artificial variables y_1 and y_2 are introduced. Since $\boldsymbol{A}_1 = \boldsymbol{e}_1$, we can drop the column associated with y_1:

$$\begin{array}{lrcrcrcrcrl} \text{minimize} & y_1 & + & y_2 & & & & & & & \\ \text{subject to} & x_1 & + & x_2 & & & + & y_1 & & & = 4, \\ & & & x_2 & - & x_3 & & & + & y_2 & = 1, \\ & x_1, & & x_2, & & x_3, & & y_1, & & y_2 & \geq 0. \end{array}$$

		x_1^*	x_2	x_3	y_2
	-5	-1	-2	1	0
$y_1^\dagger=$	4	1	1	0	0
$y_2=$	1	0	1	-1	1

Two simplex iterations then lead us to a basic feasible solution of the given problem:

		x_1	x_2^*	x_3	y_2
	-1	0	-1	1	0
$x_1=$	4	1	1	0	0
$y_2^\dagger=$	1	0	1	-1	1

		x_1	x_2	x_3
	0	0	0	0
$x_1=$	3	1	0	1
$x_2=$	1	0	1	-1

Phase II:

In the solution to the auxiliary problem, we may simply drop the artificial variable y_2 since it is not in the basis. The reduced costs of the basic variables x_1 and x_2 are zero. The reduced cost of the nonbasic variable x_3 is $c_3 - \boldsymbol{c}_B^\intercal \boldsymbol{B}^{-1}\boldsymbol{A}_3 = -1 - (1,0)\cdot(1,-1) = -2$:

		x_1	x_2	x_3^*
	-3	0	0	-2
$x_1^\dagger=$	3	1	0	1
$x_2=$	1	0	1	-1

A single simplex iteration leads to an optimal solution $(0, 4, 3)$, with optimal cost -3:

		x_1	x_2	x_3
	3	2	0	0
$x_3=$	3	1	0	1
$x_2=$	4	1	1	0

3.G. Application to transportation problems

Consider the following transportation problem. We want to ship identical items from M warehouses to N stores. Warehouse i has a_i units available and store j requires b_j units. If the total supply $\sum_{i=1}^{M} a_i$ equals the total demand $\sum_{j=1}^{N} b_j$, and the cost to ship an item from warehouse i to store j is c_{ij}, how many items x_{ij} should we ship from warehouse i to store j in order to minimize the overall transportation cost?

Problem 3.10: Given $M = 3$ warehouses and $N = 2$ stores, it is convenient to relabel $(x_{11}, x_{12}, x_{21}, x_{22}, x_{31}, x_{32})$ as $(x_1, x_2, x_3, x_4, x_5, x_6)$. Suppose the corresponding cost vector $(c_1, c_2, c_3, c_4, c_5, c_6)$ is $(1, 5, 1, 3, 1, 4)$ (in dollars), with $a_1 = 20$, $a_2 = 3$, $a_3 = 12$, $b_1 = 15$, $b_2 = 20$. How many units should each warehouse ship to each store to minimize the total transportation cost?

Since demand equals supply, each warehouse must ship all available units. The total outflow from warehouse i is $\sum_{j=1}^{2} x_{ij} = a_i$ and the total inflow into store j is $\sum_{i=1}^{3} x_{ij} = b_j$. This leads to the linear programming problem

$$
\begin{array}{lrcrcrcrcrcrcl}
\text{minimize} & x_1 & + & 5x_2 & + & x_3 & + & 3x_4 & + & x_5 & + & 4x_6 & & \\
\text{subject to} & x_1 & + & x_2 & & & & & & & & & = & 20, \\
 & & & & & x_3 & + & x_4 & & & & & = & 3, \\
 & & & & & & & & & x_5 & + & x_6 & = & 12, \\
 & x_1 & & & + & x_3 & & & + & x_5 & & & = & 15, \\
 & & & x_2 & & & + & x_4 & & & + & x_6 & = & 20, \\
 & x_1, & & x_2, & & x_3, & & x_4, & & x_5, & & x_6 & \geq & 0.
\end{array}
$$

Using the two-phase simplex method, the problem can be solved as follows.

Phase I:

Step 1. Multiply every constraint with a negative right-hand side by -1, so that $\boldsymbol{b} \geq \mathbf{0}$. In this case, the constraints are already in this form.

Step 2. In order to find a basic feasible solution, the artificial variables y_1, y_2, y_3, y_4, and y_5 are introduced to form the auxiliary problem

$$
\begin{array}{lllllllllllll}
\text{minimize} & y_1 + y_2 + y_3 + y_4 + y_5 \\
\text{subject to} & x_1 + x_2 & & & & & + y_1 & & & & & = 20, \\
 & & x_3 + x_4 & & & & & + y_2 & & & & = 3, \\
 & & & x_5 + x_6 & & & & & + y_3 & & & = 12, \\
 & x_1 & + x_3 & + x_5 & & & & & & + y_4 & & = 15, \\
 & x_2 & + x_4 & + x_6 & & & & & & & + y_5 & = 20, \\
 & x_1, & x_2, & x_3, & x_4, & x_5, & x_6, & y_1, & y_2, & y_3, & y_4, & y_5 \geq 0.
\end{array}
$$

		x_1^*	x_2	x_3	x_4	x_5	x_6	y_1	y_2	y_3	y_4	y_5
	-70	-2	-2	-2	-2	-2	-2	0	0	0	0	0
$y_1=$	20	1	1	0	0	0	0	1	0	0	0	0
$y_2=$	3	0	0	1	1	0	0	0	1	0	0	0
$y_3=$	12	0	0	0	0	1	1	0	0	1	0	0
$y_4^\dagger=$	15	1	0	1	0	1	0	0	0	0	1	0
$y_5=$	20	0	1	0	1	0	1	0	0	0	0	1

		x_1	x_2^*	x_3	x_4	x_5	x_6	y_1	y_2	y_3	y_4	y_5
	-40	0	-2	0	-2	0	-2	0	0	0	2	0
$y_1^\dagger=$	5	0	1	-1	0	-1	0	1	0	0	-1	0
$y_2=$	3	0	0	1	1	0	0	0	1	0	0	0
$y_3=$	12	0	0	0	0	1	1	0	0	1	0	0
$x_1=$	15	1	0	1	0	1	0	0	0	0	1	0
$y_5=$	20	0	1	0	1	0	1	0	0	0	0	1

		x_1	x_2	x_3^*	x_4	x_5	x_6	y_1	y_2	y_3	y_4	y_5
	-30	0	0	-2	-2	-2	-2	2	0	0	0	0
$x_2=$	5	0	1	-1	0	-1	0	1	0	0	-1	0
$y_2^\dagger=$	3	0	0	1	1	0	0	0	1	0	0	0
$y_3=$	12	0	0	0	0	1	1	0	0	1	0	0
$x_1=$	15	1	0	1	0	1	0	0	0	0	1	0
$y_5=$	15	0	0	1	1	1	1	-1	0	0	1	1

		x_1	x_2	x_3	x_4	x_5^*	x_6	y_1	y_2	y_3	y_4	y_5
	-24	0	0	0	0	-2	-2	2	2	0	0	0
$x_2=$	8	0	1	0	1	-1	0	1	1	0	-1	0
$x_3=$	3	0	0	1	1	0	0	0	1	0	0	0
$y_3=$	12	0	0	0	0	1	1	0	0	1	0	0
$x_1^\dagger=$	12	1	0	0	-1	1	0	0	-1	0	1	0
$y_5=$	12	0	0	0	0	1	1	-1	-1	0	1	1

	x_1^*	x_2	x_3	x_4	x_5	x_6
0	2	0	0	-2	0	-2
$x_2=$ 20	1	1	0	0	0	0
$x_3=$ 3	0	0	1	1	0	0
$y_3^\dagger=$ 0	-1	0	0	1	0	1
$x_5=$ 12	1	0	0	-1	1	0
$y_5=$ 0	-1	0	0	1	0	1

	x_1	x_2	x_3	x_4	x_5	x_6
0	0	0	0	0	0	0
$x_2=$ 20	0	1	0	1	0	1
$x_3=$ 3	0	0	1	1	0	0
$x_1=$ 0	1	0	0	-1	0	-1
$x_5=$ 12	0	0	0	0	1	1
$y_5=$ 0	0	0	0	0	0	0

Phase II:

	x_1	x_2	x_3	x_4^*	x_5	x_6
-115	0	0	0	-2	0	-1
$x_2=$ 20	0	1	0	1	0	1
$x_3^\dagger=$ 3	0	0	1	1	0	0
$x_1=$ 0	1	0	0	-1	0	-1
$x_5=$ 12	0	0	0	0	1	1

	x_1	x_2	x_3	x_4	x_5	x_6^*
-109	0	0	2	0	0	-1
$x_2=$ 17	0	1	-1	0	0	1
$x_4=$ 3	0	0	1	1	0	0
$x_1=$ 3	1	0	1	0	0	-1
$x_5^\dagger=$ 12	0	0	0	0	1	1

	x_1	x_2	x_3	x_4	x_5	x_6
-97	0	0	2	0	1	0
$x_2=$ 5	0	1	-1	0	-1	0
$x_4=$ 3	0	0	1	1	0	0
$x_1=$ 15	1	0	1	0	1	0
$x_6=$ 12	0	0	0	0	1	1

The minimal transportation cost of \$97 is realized by the optimal solution $(15, 5, 0, 3, 0, 12)$, where Warehouse 1 ships 15 units to Store 1 and 5 units to Store 2 and Warehouse 2 and 3 ship 3 and 12 units to Store 2, respectively.

Chapter 4

Duality

4.A. Introduction

Suppose that (x_1, x_2, x_3) is a feasible solution to the linear programming problem

$$\begin{array}{lrcrcrcl}
\text{minimize} & 4x_1 & + & 2x_2 & + & x_3 & & \\
\text{subject to} & x_1 & - & x_2 & & & \geq & 3, \\
& 2x_1 & + & x_2 & + & x_3 & \geq & 4, \\
& & & & & x_1, x_2, x_3 \geq 0. & &
\end{array}$$

If the problem has an optimal cost, we can find a lower bound for the optimal cost with the following procedure. Let p_1 and p_2 be nonnegative numbers. On summing the first constraint multiplied by p_1 with the second constraint multiplied by p_2, we obtain

$$p_1(x_1 - x_2) + p_2(2x_1 + x_2 + x_3) \geq 3p_1 + 4p_2.$$

On rearranging this result, we find

$$x_1(p_1 + 2p_2) + x_2(-p_1 + p_2) + x_3 p_2 \geq 3p_1 + 4p_2.$$

If we enforce the constraints $p_1 + 2p_2 \leq 4$, $-p_1 + p_2 \leq 2$, and $p_2 \leq 1$, we obtain a lower bound to the optimal cost:

$$4x_1 + 2x_2 + x_3 \geq x_1(p_1 + 2p_2) + x_2(-p_1 + p_2) + x_3 p_2 \geq 3p_1 + 4p_2.$$

The linear programming problem

$$\begin{array}{llr}
\text{maximize} & 3p_1 + 4p_2 & \\
\text{subject to} & p_1 + 2p_2 & \leq 4, \\
& -p_1 + p_2 & \leq 2, \\
& p_2 & \leq 1, \\
& p_1, p_2 & \geq 0
\end{array}$$

that determines the **largest possible value** for our lower bound to the optimal cost is known as the *dual problem* to the original linear programming problem.

Remark: Duality theory can be also motivated as an outgrowth of the Lagrange multiplier method.

- Consider the optimization problem

$$\begin{array}{lcccc} \text{minimize} & x^2 & + & y^2 & \\ \text{subject to} & x & + & y & = & 1. \end{array}$$

Instead of enforcing the constraint $x+y=1$, we allow it to be violated and associate a constant Lagrange multiplier, or *price*, p with the amount $1-x-y$ by which the constraint is violated. Instead of minimizing the cost subject to the constraint $x+y=1$, we minimize the Lagrangian

$$L(x,y,p) = x^2 + y^2 + p(1-x-y)$$

over all x and y, without further constraints:

$$0 = \frac{\partial L}{\partial x} = 2x - p,$$

$$0 = \frac{\partial L}{\partial y} = 2y - p.$$

The Lagrangian takes on its minimum value at the critical point $(x,y) = (p/2, p/2)$, which depends on p. If we now enforce the constraint $x+y=1$, we see that $p=1$ and $x=y=1/2$. At this value of p, the presence or absence of the constraint does not affect L: an optimal solution of the unconstrained problem is thus also an optimal solution of the original constrained problem.

In a similar manner, we can associate a price variable with each constraint of a linear programming problem and search for prices under which the presence or absence of the constraints does not affect the optimal cost. The right prices can be found by solving a new linear programming problem that is the *dual* of the original problem.

Consider the standard-form problem

$$\begin{array}{ll} \text{minimize} & \boldsymbol{c}^\mathsf{T}\boldsymbol{x} \\ \text{subject to} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \\ & \boldsymbol{x} \geq \boldsymbol{0}, \end{array}$$

which we will call the *primal problem*, and assume that an optimal solution $\boldsymbol{x}^*$ exists. Let us introduce a *relaxed problem* in which the constraint $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ is replaced by a penalty $\boldsymbol{p}^\mathsf{T}(\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x})$, where $\boldsymbol{p}$ is a price vector with the same length as $\boldsymbol{b}$. This results in the following problem

$$\begin{array}{ll} \text{minimize} & \boldsymbol{c}^\mathsf{T}\boldsymbol{x} + \boldsymbol{p}^\mathsf{T}(\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}) \\ \text{subject to} & \boldsymbol{x} \geq \boldsymbol{0}. \end{array}$$

Let $g(\boldsymbol{p})$ be the optimal cost for the relaxed problem. We expect $g(\boldsymbol{p})$ to be no larger than the optimal primal cost $\boldsymbol{c}^\mathsf{T}\boldsymbol{x}^*$:

$$g(\boldsymbol{p}) = \min_{\boldsymbol{x}\geq 0} [\boldsymbol{c}^\mathsf{T}\boldsymbol{x} + \boldsymbol{p}^\mathsf{T}(\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x})] \leq \boldsymbol{c}^\mathsf{T}\boldsymbol{x}^* + \boldsymbol{p}^\mathsf{T}(\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}^*) = \boldsymbol{c}^\mathsf{T}\boldsymbol{x}^*,$$

noting that $\boldsymbol{A}\boldsymbol{x}^* = \boldsymbol{b}$. Thus, for every price vector $\boldsymbol{p}$, the value of $g(\boldsymbol{p})$ provides a lower bound to the optimal cost $\boldsymbol{c}^\mathsf{T}\boldsymbol{x}^*$. The tightest possible lower bound can be found by solving the problem

$$\begin{array}{ll} \text{maximize} & g(\boldsymbol{p}) \\ \text{subject to} & \text{no constraints,} \end{array}$$

which is known as the *dual* problem. How close is this tightest lower bound to the actual optimal value $\boldsymbol{c}^\mathsf{T}\boldsymbol{x}^*$ of the primal problem? As we will see, an important result of duality theory asserts that the optimal value $g(\boldsymbol{p})$ of the dual problem in fact equals the optimal value $\boldsymbol{c}^\mathsf{T}\boldsymbol{x}^*$ of the primal problem. In other words, when the price vector $\boldsymbol{p}$ is chosen to optimize the dual problem, the option of violating the constraint $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ is of no value. Note that

$$g(\boldsymbol{p}) = \min_{\boldsymbol{x}\geq \boldsymbol{0}} [\boldsymbol{c}^\mathsf{T}\boldsymbol{x} + \boldsymbol{p}^\mathsf{T}(\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x})] = \boldsymbol{p}^\mathsf{T}\boldsymbol{b} + \min_{\boldsymbol{x}\geq \boldsymbol{0}} (\boldsymbol{c}^\mathsf{T} - \boldsymbol{p}^\mathsf{T}\boldsymbol{A})\boldsymbol{x},$$

where

$$\min_{\boldsymbol{x}\geq \boldsymbol{0}} (\boldsymbol{c}^\mathsf{T} - \boldsymbol{p}^\mathsf{T}\boldsymbol{A})\boldsymbol{x} = \begin{cases} 0 & \text{if } \boldsymbol{c}^\mathsf{T} - \boldsymbol{p}^\mathsf{T}\boldsymbol{A} \geq \boldsymbol{0}^\mathsf{T}, \\ -\infty & \text{otherwise.} \end{cases}$$

In maximizing $g(\boldsymbol{p})$, we only need to consider those values of $\boldsymbol{p}$ for which $g(\boldsymbol{p})$ is not $-\infty$. Hence, the dual problem is equivalent to the linear programming problem

$$\begin{array}{ll} \text{maximize} & \boldsymbol{p}^\mathsf{T}\boldsymbol{b} \\ \text{subject to} & \boldsymbol{p}^\mathsf{T}\boldsymbol{A} \leq \boldsymbol{c}^\mathsf{T}. \end{array}$$

Remark: If instead of the equality constraint $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$, the primal problem had inequality constraints of the form $\boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b}$, they could be replaced by $\boldsymbol{A}\boldsymbol{x} - \boldsymbol{s} = \boldsymbol{b}$ and $\boldsymbol{s} \geq \boldsymbol{0}$, with cost function $\boldsymbol{c}^\mathsf{T}\boldsymbol{x} + \boldsymbol{0}^\mathsf{T}\boldsymbol{s}$:

$$[\boldsymbol{A}|{-\boldsymbol{1}}]\begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{s} \end{bmatrix} = \boldsymbol{b},$$

leading to the dual constraints

$$\boldsymbol{p}^\mathsf{T}[\boldsymbol{A}|{-\boldsymbol{1}}] \leq [\boldsymbol{c}^\mathsf{T}|\boldsymbol{0}^\mathsf{T}];$$

that is,

$$\boldsymbol{p}^\mathsf{T}\boldsymbol{A} \leq \boldsymbol{c}^\mathsf{T}, \qquad \boldsymbol{p} \geq \boldsymbol{0}.$$

Remark: Alternatively, if $\boldsymbol{x}$ were a free variable, we would use the fact that

$$\min_{\boldsymbol{x}}(\boldsymbol{c}^\intercal - \boldsymbol{p}^\intercal \boldsymbol{A})\boldsymbol{x} = \begin{cases} 0 & \text{if } \boldsymbol{c}^\intercal - \boldsymbol{p}^\intercal \boldsymbol{A} = \boldsymbol{0}^\intercal, \\ -\infty & \text{otherwise,} \end{cases}$$

leading to the constraint $\boldsymbol{p}^\intercal \boldsymbol{A} = \boldsymbol{c}^\intercal$ in the dual problem.

4.B. The dual problem

Let $\boldsymbol{A}$ be a matrix with rows $\boldsymbol{a}_i^\intercal$ and columns $\boldsymbol{A}_j$. The objective function $\boldsymbol{p}^\intercal \boldsymbol{b}$ in the dual problem can be equivalently written as $\boldsymbol{b}^\intercal \boldsymbol{p}$. Moreover, by taking a transpose, dual row constraints like $\boldsymbol{p}^\intercal \boldsymbol{A} = \boldsymbol{c}^\intercal$ (and inequality versions thereof) can be rewritten as column constraints like $\boldsymbol{A}^\intercal \boldsymbol{p} = \boldsymbol{c}$. One then notices that the transformation between the primal and the dual problems can be accomplished by swapping $\boldsymbol{x}$ and $\boldsymbol{p}$, swapping $\boldsymbol{b}$ and $\boldsymbol{c}$, and taking the transpose of $\boldsymbol{A}$. For a *primal* minimization problem with the structure shown on the left, the corresponding *dual* maximization problem is listed on the right:

$$\begin{array}{llll} \text{minimize} & \boldsymbol{c}^\intercal \boldsymbol{x} & & \\ \text{subject to} & \boldsymbol{a}_i^\intercal \boldsymbol{x} \le b_i, & i \in \mathcal{M}_1, \\ & \boldsymbol{a}_i^\intercal \boldsymbol{x} \ge b_i, & i \in \mathcal{M}_2, \\ & \boldsymbol{a}_i^\intercal \boldsymbol{x} = b_i, & i \in \mathcal{M}_3, \\ & x_j \le 0, & j \in \mathcal{N}_1, \\ & x_j \ge 0, & j \in \mathcal{N}_2, \\ & x_j \text{ free}, & j \in \mathcal{N}_3, \end{array} \qquad \begin{array}{lll} \text{maximize} & \boldsymbol{b}^\intercal \boldsymbol{p} & \\ \text{subject to} & p_i \le 0, & i \in \mathcal{M}_1, \\ & p_i \ge 0, & i \in \mathcal{M}_2, \\ & p_i \text{ free}, & i \in \mathcal{M}_3, \\ & \boldsymbol{A}_j^\intercal \boldsymbol{p} \ge c_j, & j \in \mathcal{N}_1, \\ & \boldsymbol{A}_j^\intercal \boldsymbol{p} \le c_j, & j \in \mathcal{N}_2, \\ & \boldsymbol{A}_j^\intercal \boldsymbol{p} = c_j, & j \in \mathcal{N}_3. \end{array} \tag{4.1}$$

Remark: Each (inequality or equality) constraint in the primal problem is associated with a (sign-constrained or free) variable in the dual problem, and vice-versa:

	minimize	maximize	
constraints	$\le b_i$ $\ge b_i$ $= b_i$	≤ 0 ≥ 0 free	variables
variables	≤ 0 ≥ 0 free	$\ge c_j$ $\le c_j$ $= c_j$	constraints

- For example, the dual to the linear programming problem

$$\begin{array}{lrcrcrl}
\text{minimize} & x_1 & + & 2x_2 & + & 3x_3 & \\
\text{subject to} & -x_1 & + & 3x_2 & & & = 5, \\
& 2x_1 & - & x_2 & + & 3x_3 & \geq 6, \\
& & & & & x_3 & \leq 4, \\
& x_1 \geq 0, & & & & & \\
& x_2 \leq 0. & & & & &
\end{array}$$

is

$$\begin{array}{lrcrcrl}
\text{maximize} & 5p_1 & + & 6p_2 & + & 4p_3 & \\
\text{subject to} & -p_1 & + & 2p_2 & & & \leq 1, \\
& 3p_1 & - & p_2 & & & \geq 2, \\
& & & 3p_2 & + & p_3 & = 3, \\
& p_2 \geq 0, & & & & & \\
& p_3 \leq 0. & & & & &
\end{array}$$

- If we rename each p_i in the dual problem of the previous example to x_i, we obtain a linear programming problem equivalent to this minimization problem:

$$\begin{array}{lrcrcrl}
\text{minimize} & -5x_1 & - & 6x_2 & - & 4x_3 & \\
\text{subject to} & x_1 & - & 2x_2 & & & \geq -1, \\
& -3x_1 & + & x_2 & & & \leq -2, \\
& & & -3x_2 & - & x_3 & = -3, \\
& x_2 \geq 0, & & & & & \\
& x_3 \leq 0. & & & & &
\end{array}$$

We can then formulate the dual to the above minimization problem:

$$\begin{array}{lrcrcrl}
\text{maximize} & -p_1 & - & 2p_2 & - & 3p_3 & \\
\text{subject to} & p_1 & - & 3p_2 & & & = -5, \\
& -2p_1 & + & p_2 & - & 3p_3 & \leq -6, \\
& & & & & -p_3 & \geq -4, \\
& p_1 \geq 0, & & & & & \\
& p_2 \leq 0. & & & & &
\end{array}$$

Notice that we have arrived back at the equivalent maximization version of the primal problem!

Remark: The following theorem expresses the general behaviour we observed in the previous example: the dual to the dual is equivalent to the primal problem.

Theorem 4.1: *If we transform the dual into an equivalent minimization problem and then form its dual, we obtain a problem equivalent to the original problem.*

Remark: The following result is also straightforward to show.

Theorem 4.2: *Suppose that a linear programming problem P is transformed to another linear programming problem P' by a sequence of transformations of the following types:*

1. *Replace each free variable with the difference of two nonnegative variables.*
2. *Replace each inequality constraint by an equality constraint involving a nonnegative slack variable.*
3. *If some row of the resulting standard-form matrix $\boldsymbol{A}$ is a linear combination of the other rows, eliminate the corresponding equality constraint.*

Then the duals of P and P' are equivalent in the sense that they are either both infeasible or have the same optimal cost.

4.C. The duality theorem

Remark: Notice from (4.1) that if $\boldsymbol{x}$ and $\boldsymbol{p}$ are feasible solutions to the primal and dual problems, respectively, $u_i = p_i(\boldsymbol{a}_i^\top \boldsymbol{x} - b_i) \geq 0$ for $i = 1, \ldots, m$ and $v_j = (c_j - \boldsymbol{A}_j^\top \boldsymbol{p})x_j \geq 0$ for $j = 1, \ldots, n$. Furthermore,

$$0 \leq \sum_{i=1}^{m} u_i + \sum_{j=1}^{n} v_j = \boldsymbol{p}^\top(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}) + (\boldsymbol{c} - \boldsymbol{A}^\top \boldsymbol{p})^\top \boldsymbol{x} = \boldsymbol{p}^\top(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}) + (\boldsymbol{c}^\top - \boldsymbol{p}^\top \boldsymbol{A})\boldsymbol{x} = -\boldsymbol{p}^\top \boldsymbol{b} + \boldsymbol{c}^\top \boldsymbol{x}.$$

This inequality is expressed in the following theorem.

Theorem 4.3 (Weak duality): *If $\boldsymbol{x}$ is a feasible solution to the primal problem and $\boldsymbol{p}$ is a feasible solution to the dual problem, then*

$$\boldsymbol{p}^\top \boldsymbol{b} \leq \boldsymbol{c}^\top \boldsymbol{x}.$$

Corollary 4.3.1:

(i) If the optimal cost in the primal problem is $-\infty$, the dual problem is infeasible.

(ii) If the optimal cost in the dual problem is ∞, the primal problem is infeasible.

Proof:
(i) Suppose that the optimal cost in the primal problem is $-\infty$; this means that there are feasible solutions $\boldsymbol{x}$ for which the cost can be made as small as desired. If the dual problem had a feasible solution $\boldsymbol{p}$, then by Theorem 4.3 we would know that $\boldsymbol{p}^\top \boldsymbol{b} \leq \boldsymbol{c}^\top \boldsymbol{x}$ for every feasible solution $\boldsymbol{x}$ of the primal problem. The existence of such a lower bound on the optimal cost of the primal problem would contradict the premise that its optimal cost is $-\infty$. Thus, the dual problem must in fact be infeasible.

(ii) Exercise.

Corollary 4.3.2: Let $\boldsymbol{x}$ and $\boldsymbol{p}$ be feasible solutions to the primal and dual problems, respectively, and suppose that $\boldsymbol{p}^\intercal \boldsymbol{b} = \boldsymbol{c}^\intercal \boldsymbol{x}$. Then $\boldsymbol{x}$ and $\boldsymbol{p}$ are optimal solutions.

Proof: For every feasible primal solution $\boldsymbol{y}$ and feasible dual solution $\boldsymbol{q}$, Theorem 4.3 guarantees that $\boldsymbol{q}^\intercal \boldsymbol{b} \le \boldsymbol{c}^\intercal \boldsymbol{x} = \boldsymbol{p}^\intercal \boldsymbol{b} \le \boldsymbol{c}^\intercal \boldsymbol{y}$, from which we see that $\boldsymbol{x}$ and $\boldsymbol{p}$ are optimal.

Theorem 4.4 (Strong duality): *If a linear programming problem has an optimal solution, so does its dual, and the respective optimal costs are equal.*

Proof: Given a linear programming problem Π' with an optimal solution, transform the linear programming into an equivalent standard-form problem Π with the same optimal cost, such that the equality constraint matrix $\boldsymbol{A}$ has linearly independent rows. Suppose that Π has the form

$$\begin{array}{ll} \text{minimize} & \boldsymbol{c}^\intercal \boldsymbol{x} \\ \text{subject to} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \\ & \quad \boldsymbol{x} \ge \boldsymbol{0}. \end{array}$$

Using Bland's rule, the simplex method terminates with an optimal solution $\boldsymbol{x}$, optimal basis $\boldsymbol{B}$, and reduced cost vector

$$\bar{\boldsymbol{c}}^\intercal = \boldsymbol{c}^\intercal - \boldsymbol{p}^\intercal \boldsymbol{A} \ge \boldsymbol{0},$$

where $\boldsymbol{p}^\intercal = \boldsymbol{c}_B^\intercal \boldsymbol{B}^{-1}$ is defined in terms of the basic cost vector $\boldsymbol{c}_B$. Then $\boldsymbol{p}^\intercal \boldsymbol{A} \le \boldsymbol{c}^\intercal$, so that $\boldsymbol{p}$ is a feasible solution to the dual problem

$$\begin{array}{ll} \text{maximize} & \boldsymbol{p}^\intercal \boldsymbol{b} \\ \text{subject to} & \boldsymbol{A}^T \boldsymbol{p} \le \boldsymbol{c}. \end{array}$$

Furthermore

$$\boldsymbol{p}^\intercal \boldsymbol{b} = \boldsymbol{c}_B^\intercal \boldsymbol{B}^{-1} \boldsymbol{b} = \boldsymbol{c}_B^\intercal \boldsymbol{x}_B = \boldsymbol{c}^\intercal \boldsymbol{x},$$

where $\boldsymbol{x}_B = \boldsymbol{B}^{-1}\boldsymbol{b}$ is the vector of basic variables. Corollary 4.3.2 then tells us that $\boldsymbol{p}$ is an optimal solution to the dual problem, with optimal cost $\boldsymbol{p}^\intercal \boldsymbol{b}$ equal to the optimal primal cost $\boldsymbol{c}^\intercal \boldsymbol{x}$. That is, the dual of Π has the same optimal cost as Π. Theorem 4.2 guarantees that the duals of Π and Π' have identical optimal costs, both equal to the optimal primal cost of Π, which has the same optimal cost as the original linear programming problem Π'.

Remark: It is possible for both the primal and dual problems to be infeasible. The infeasible primal problem

$$\begin{array}{lcccccl} \text{minimize} & x_1 & + & 2x_2 & & \\ \text{subject to} & x_1 & + & x_2 & = & 1, \\ & 2x_1 & + & 2x_2 & = & 3 \end{array}$$

has dual

$$\begin{array}{lrcrcl} \text{maximize} & p_1 & + & 3p_2 & & \\ \text{subject to} & p_1 & + & 2p_2 & = & 1 \\ & p_1 & + & 2p_2 & = & 2, \end{array}$$

which is also infeasible.

Remark: Recall that there are three possible outcomes when solving a linear programming problem:

1. There exists an optimal solution;
2. The problem is *unbounded*: the optimal cost is $-\infty$ (∞) for minimization (maximation) problems;
3. The problem is infeasible.

This leads to nine combinations of outcomes for the primal and dual problems. Theorem 4.4 guarantees that if one problem has an optimal solution, so does the other. Moreover, Corollary 4.3.1 establishes that if one problem is unbounded, the other problem is infeasible. This allows us to complete the following table describing the various possible outcomes for a primal problem and its dual. We observe that of the nine outcomes, only four are possible.

Primal \ Dual	Finite optimum	Unbounded	Infeasible
Finite optimum	Possible	Impossible	Impossible
Unbounded	Impossible	Impossible	Possible
Infeasible	Impossible	Possible	Possible

4.D. Complementary slackness

An important relationship between the primal and dual optimal solutions is provided by the *complementary slackness* conditions:

Theorem 4.5 (Complementary slackness): *Let $\boldsymbol{x}$ and $\boldsymbol{p}$ be feasible solutions to the primal and dual problems, respectively. Then $\boldsymbol{x}$ and $\boldsymbol{p}$ are optimal solutions if and only if*

$$p_i(\boldsymbol{a}_i^\intercal \boldsymbol{x} - b_i) = 0, \quad \textit{for all } i = 1, \ldots, m$$

and

$$(c_j - \boldsymbol{A}_j^\intercal \boldsymbol{p})x_j = 0, \quad \textit{for all } j = 1, \ldots, n.$$

Proof: We previously noted for feasible solutions $\boldsymbol{x}$ and $\boldsymbol{p}$ that $u_i = p_i(\boldsymbol{a}_i^\mathsf{T}\boldsymbol{x} - b_i) \geq 0$ for each $i = 1, \ldots, m$ and $v_j = (c_j - \boldsymbol{A}_j^\mathsf{T}\boldsymbol{p})x_j \geq 0$ for $j = 1, \ldots n$. Moreover, we found

$$0 \leq \sum_{i=1}^{m} u_i + \sum_{j=1}^{n} v_j = -\boldsymbol{p}^\mathsf{T}\boldsymbol{b} + \boldsymbol{c}^\mathsf{T}\boldsymbol{x}.$$

If $\boldsymbol{x}$ and $\boldsymbol{p}$ are optimal then Theorem 4.4 implies that the right-hand side is zero, from which the desired conditions immediately follow. The converse result follows on applying Corollary 4.3.2.

Remark: Complementary slackness means that $\boldsymbol{x}$ and $\boldsymbol{p}$ are optimal iff for each i either $\boldsymbol{a}_i^\mathsf{T}\boldsymbol{x} = b_i$ or $p_i = 0$ and for each j either $\boldsymbol{A}_j^\mathsf{T}\boldsymbol{p} = c_j$ or $x_j = 0$. That is, in each case there must be an active constraint in one of the two domains.

Problem 4.1: Determine if $\boldsymbol{x} = (2/5, 1/5, 0)$ is an optimal solution to the problem

$$\begin{array}{lrcrcrcr}
\text{minimize} & -3x_1 & - & 6x_2 & - & 2x_3 & & \\
\text{subject to} & -3x_1 & - & 4x_2 & - & x_3 & \geq & -2, \\
 & -x_1 & - & 3x_2 & - & 2x_3 & \geq & -1, \\
 & & & x_1, x_2, x_3 \geq 0. & & & &
\end{array}$$

The dual problem is

$$\begin{array}{lrcrcr}
\text{maximize} & -2p_1 & - & p_2 & & \\
\text{subject to} & -3p_1 & - & p_2 & \leq & -3, \\
 & -4p_1 & - & 3p_2 & \leq & -6, \\
 & -p_1 & - & 2p_2 & \leq & -2, \\
 & & & p_1, p_2 \geq 0. & &
\end{array}$$

For the solution $\boldsymbol{x} = (2/5, 1/5, 0)$, since

$$\begin{aligned}
-3x_1 - 4x_2 - x_3 &= -2, \\
-x_1 - 3x_2 - 2x_3 &= -1,
\end{aligned}$$

we see that $\boldsymbol{x}$ is a feasible solution, but because these two constraints are active, we obtain no information on p_1 and p_2. Since $x_1 > 0$ and $x_2 > 0$ but $x_3 = 0$, optimality also requires

$$\begin{aligned}
-3p_1 - p_2 &= -3, \\
-4p_1 - 3p_2 &= -6,
\end{aligned}$$

which has the unique solution $p_1 = 3/5$, $p_2 = 6/5$. Since $-p_1 - 2p_2 = -3 \leq -2$, we see that this represents a feasible solution to the dual problem. Since both $\boldsymbol{x}$ and $\boldsymbol{p}$ are feasible solutions that satisfy all of the complementary slackness conditions, we conclude that these are optimal solutions. Note that the optimal cost for both problems is $-12/5$.

Problem 4.2: Determine if $\boldsymbol{x} = (0, 2)$ is an optimal solution to the problem

$$\begin{array}{lrcrcr} \text{minimize} & 3x_1 & + & x_2 & & \\ \text{subject to} & 2x_1 & + & x_2 & \geq & 2, \\ & 3x_1 & + & 5x_2 & \geq & 10, \\ & x_1, x_2 \geq 0. & & & & \end{array}$$

Consider the dual to the given problem:

$$\begin{array}{lrcrcr} \text{maximize} & 2p_1 & + & 10p_2 & & \\ \text{subject to} & 2p_1 & + & 3p_2 & \leq & 3, \\ & p_1 & + & 5p_2 & \leq & 1, \\ & p_1, p_2 \geq 0. & & & & \end{array}$$

At the given $\boldsymbol{x}$, we see that the first two primal constraints are active. Since $x_2 \neq 0$, complementary slackness requires that $p_1 + 5p_2 = 1$, which means that $p_1 = 1 - 5p_2$. For $p_2 \geq 0$, the first constraint is automatically satisfied: $2(1 - 5p_2) + 3p_2 = 2 - 7p_2 \leq 2 < 3$. Furthermore, $p_1 \geq 0$ requires that $p_2 \leq 1/5$. Any solution of the form $(1 - 5p_2, p_2)$ where $p_2 \in [0, 1/5]$ is a feasible solution of the dual problem that satisfies the complementary slackness conditions and is thus an optimal solution, with optimal cost 2. Likewise, $\boldsymbol{x} = (0, 2)$ is an optimal solution to the primal problem, also with optimal cost 2.

4.E. Farkas' lemma and linear inequalities

Consider the set of standard-form constraints $\boldsymbol{Ax} = \boldsymbol{b}, \boldsymbol{x} \geq \boldsymbol{0}$. If there exists some $\boldsymbol{p}$ such that $\boldsymbol{p}^\mathsf{T}\boldsymbol{A} \geq \boldsymbol{0}^\mathsf{T}$ and $\boldsymbol{p}^\mathsf{T}\boldsymbol{b} < 0$, then $\boldsymbol{x} \geq \boldsymbol{0} \Rightarrow \boldsymbol{p}^\mathsf{T}\boldsymbol{Ax} \geq 0$, from which we see that $\boldsymbol{Ax} \neq \boldsymbol{b}$. We have thus found a *certificate of infeasibility* for this linear programming problem. This is expressed in the following theorem.

Theorem 4.6 (Farkas' lemma): *Let $\boldsymbol{A}$ be an $m \times n$ matrix and let $\boldsymbol{b} \in \mathbb{R}^m$. Exactly one of the following alternatives holds:*

(i) *$\boldsymbol{Ax} = \boldsymbol{b}$ has a solution $\boldsymbol{x} \geq \boldsymbol{0}$;*

(ii) *$\boldsymbol{A}^\mathsf{T}\boldsymbol{p} \geq \boldsymbol{0}$ has a solution $\boldsymbol{p}$ with $\boldsymbol{p}^\mathsf{T}\boldsymbol{b} < 0$.*

Proof. One direction is easy. If there exists some $\boldsymbol{x} \geq \boldsymbol{0}$ satisfying $\boldsymbol{Ax} = \boldsymbol{b}$ and if $\boldsymbol{A}^\mathsf{T}\boldsymbol{p} \geq \boldsymbol{0}$, then $\boldsymbol{p}^\mathsf{T}\boldsymbol{b} = \boldsymbol{p}^\mathsf{T}\boldsymbol{Ax} = \boldsymbol{x}^\mathsf{T}\boldsymbol{A}^\mathsf{T}\boldsymbol{p} \geq 0$, so the second alternative cannot hold.

Let us now assume that there exists no vector $\boldsymbol{x} \geq \boldsymbol{0}$ satisfying $\boldsymbol{Ax} = \boldsymbol{b}$. Consider the pair of problems

$$\begin{array}{lll@{\qquad}ll} \text{maximize} & \boldsymbol{0}^\mathsf{T}\boldsymbol{x} & & \text{minimize} & \boldsymbol{p}^\mathsf{T}\boldsymbol{b} \\ \text{subject to} & \boldsymbol{Ax} = \boldsymbol{b}, & & \text{subject to} & \boldsymbol{A}^\mathsf{T}\boldsymbol{p} \geq \boldsymbol{0} \\ & \boldsymbol{x} \geq \boldsymbol{0}, & & & \end{array}$$

and note that the first is the dual of the second. We are given that the maximization problem is infeasible, which implies that the minimization problem is either unbounded (the optimal cost is $-\infty$) or infeasible. Since $\boldsymbol{p} = \boldsymbol{0}$ is a feasible solution to the minimization problem, it follows that the minimization problem is unbounded. In particular, there exists some feasible $\boldsymbol{p}$ for which $\boldsymbol{A}^\intercal \boldsymbol{p} \geq \boldsymbol{0}$ and the cost is negative: $\boldsymbol{p}^\intercal \boldsymbol{b} < 0$.

Problem 4.3: Use Farka's lemma to show that the following system of inequalities is inconsistent.

$$\begin{array}{rcrcrcr} x_1 & + & x_2 & + & x_3 & = & 1, \\ x_1 & - & x_2 & + & 2x_3 & = & -2, \\ \multicolumn{5}{l}{x_1, x_2, x_3 \geq 0.} \end{array}$$

Since $p_1 = p_2 = 1$ satisfies the dual constraints

$$\begin{array}{rcrcr} p_1 & + & p_2 & \geq & 0, \\ p_1 & - & p_2 & \geq & 0, \\ p_1 & + & 2p_2 & \geq & 0, \end{array}$$

but $p_1 - 2p_2 = -1 < 0$, we see by Farkas' lemma that the given equations are inconsistent.

Problem 4.4: Use a generalization of Farkas' lemma to determine whether the following set of inequalities is consistent.

$$\begin{array}{rcrcr} x_1 & & & \leq & 2, \\ & & x_2 & \leq & -3, \\ -x_1 & - & 2x_2 & \leq & 5, \end{array}$$

Consider the dual problems

$$\begin{array}{llrcrcr} \text{maximize} & \boldsymbol{0}^\intercal \boldsymbol{x} & & & & & \\ \text{subject to} & & x_1 & & & \leq & 2, \\ & & & & x_2 & \leq & -3, \\ & & -x_1 & - & 2x_2 & \leq & 5, \end{array} \qquad \begin{array}{llrcrcrcr} \text{minimize} & 2p_1 & - & 3p_2 & + & 5p_3 & & \\ \text{subject to} & p_1 & & & - & p_3 & = & 0, \\ & & & p_2 & - & 2p_3 & = & 0, \\ & \multicolumn{7}{l}{p_1, p_2, p_3 \geq 0.} \end{array}$$

A feasible solution of the minimization problem requires $p_1 = p_3$ and $p_2 = 2p_3$, so that the cost $\boldsymbol{p}^\intercal \boldsymbol{b}$ reduces to $2p_3 - 3(2p_3) + 5p_3 = p_3 \geq 0$. Thus the minimization problem is not unbounded. The solution $p_1 = p_2 = p_3 = 0$ satisfies the constraints, so the minimization problem is not infeasible either. That means that the minimization problem has an optimal solution (the one we just considered). The maximization problem is therefore feasible; that is, the original inequalities are consistent.

4.F. Factory and supply-chain perspective

4.F.1 A manufacturer in a supply chain

Consider a factory that produces two finished products: tables and chairs. The factory uses two scarce resources owned by upstream suppliers: wood and labour.

The resources required to manufacture each product and their availability are given in the following table, along with the values of the finished products:

Resource \ Product	**Table (x_1)**	**Chair (x_2)**	**Supply ($\boldsymbol{b}$)**
Value ($\boldsymbol{c}$)	50	30	
Wood (p_1)	3	2	100
Labour (p_2)	1	1	80

4.F.2 Primal problem: production planning

Let x_1 = number of tables produced, x_2 = number of chairs produced. The factory wants to maximize the value created from the available resources:

$$\begin{aligned} \text{maximize} \quad & 50x_1 + 30x_2 \quad \text{(value created)} \\ \text{subject to} \quad & 3x_1 + 2x_2 \leq 100 \quad \text{(wood constraint)}, \\ & x_1 + x_2 \leq 80 \quad \text{(labour constraint)}, \\ & x_1, x_2 \geq 0. \end{aligned}$$

The primal problem asks the question:

How should scarce resources be best used to generate the most value?

4.F.3 Dual problem: resource pricing (upstream view)

To understand resource pricing, we introduce the concept of *arbitrage*: earning a risk-free profit by buying low and selling high. Resource prices are driven to the *lowest levels that prevent risk-free arbitrage through production.*

If the resource prices were to fall below this level, the factory could earn an immediate margin on every unit produced; if they were to rise above this level, the factory would not be interested in buying the resources.

Let p_1 = value of one unit of wood, p_2 = value of one unit of labour. In an ideal market, the value of each finished product would never be more than the cost of its inputs:

$$50 \leq 3p_1 + p_2 \quad \text{(table)},$$

$$30 \leq 2p_1 + p_2 \quad \text{(chair)}.$$

Thus the dual problem is:

$$\begin{aligned} \text{minimize} \quad & 100p_1 + 80p_2 \quad \text{(total value of resources)} \\ \text{subject to} \quad & 3p_1 + p_2 \geq 50 \quad \text{(table no-arbitrage)}, \\ & 2p_1 + p_2 \geq 30 \quad \text{(chair no-arbitrage)}, \\ & p_1, p_2 \geq 0. \end{aligned}$$

Effectively, the dual problem assigns a market-consistent value to each resource, converting the production plan into an equivalent valuation of the underlying inputs in a competitive market.

The dual problem answers a different question than the primal problem:

> *What is the lowest set of input prices that prevents the factory from gaining a positive margin on any combination of products?*

While the primal problem determines how much to produce, the dual problem determines how valuable the input resources must be.

4.F.4 Economic meaning of strong duality

A fundamental result of linear programming is that at an equilibrium, the maximum value created by production equals the minimum value of available resources. In the language of economics, this means:

> At equilibrium, the best possible use of resources (the manufacturer's production plan) generates exactly the same value as the resources are worth when priced by a competitive market.

This is a statement of *market clearing*: the factory cannot extract more value than is embedded in the resources. The suppliers cannot charge more than the value those resources generate. No arbitrage opportunity remains for either the factory or the suppliers.

The primal problem chooses *quantities* to maximize value; the dual problem chooses resource *prices* to eliminate arbitrage: equilibrium happens when resource value and product value agree, leaving no room for arbitrage.

4.F.5 Complementary slackness: scarcity creates value

Duality also explains why only *scarce* resources are valuable. At an optimum:

- If a resource is *fully used* (constraint is active), its price may be positive.
- If a resource is *not fully used* (constraint is inactive), its price is zero.

This can be expressed in economic terms:

> *Only scarce resources command a price.*

Thus, duality in linear programming is a formalization of how value, scarcity, and equilibrium prices emerge in a competitive supply chain.

Remark: Where then does real-world profit come from? The dual constraints

$$\boldsymbol{A}_j^\intercal \boldsymbol{p} \geq c_j$$

implicitly assume that the value of a finished product cannot exceed the value of the competitive inputs required to produce it. Duality does not provide a theory of *corporate profit*. Instead, it describes the maximum value that can be attributed *solely to scarce resources* under perfect competition.

Remark: In the real world, firms often sell products for more than the value of their raw inputs due to factors outside this simplistic model:

- proprietary knowledge and trade secrets;
- patents and legal monopolies;
- superior technology and process advantages;
- branding and consumer perception;
- network effects and platform dominance;
- asymmetric information;
- organizational or logistical efficiency.

These sources of additional return are collectively captured by the *markup* of the product over its input cost. Duality in linear programming accounts only for the value explained by resource scarcity; any markup reflects firm-specific advantages and market conditions. The primal and dual problems determine the *imputed resource value* of a product, not its full market potential. A selling price above this implied value reflects innovation, market power, or structural advantage—not physical scarcity.

> *Duality calculates the hidden* shadow *value of the world as if it were perfectly fair. Profit, in the real world, is a deviation from that fairness.*

Chapter 5

Sensitivity Analysis

In this chapter, we consider the standard-form problem

$$\begin{aligned} \text{minimize} \quad & \boldsymbol{c}^\intercal \boldsymbol{x} \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \\ & \boldsymbol{x} \geq \boldsymbol{0}. \end{aligned}$$

We assume that we already have an optimal basis $\boldsymbol{B}$ and the associated optimal solution $\boldsymbol{x}^*$. We want to see what happens when some entry of $\boldsymbol{A}$, $\boldsymbol{b}$, or $\boldsymbol{c}$ is changed. We first examine under what conditions $\boldsymbol{B}$ remains optimal, while maintaining the following properties:

- feasibility: $\boldsymbol{x}_B = \boldsymbol{B}^{-1}\boldsymbol{b} \geq \boldsymbol{0}$;
- optimality: $\bar{\boldsymbol{c}}^\intercal = \boldsymbol{c}^\intercal - \boldsymbol{c}_B^\intercal \boldsymbol{B}^{-1}\boldsymbol{A} \geq \boldsymbol{0}^\intercal$.

5.A. A new non-negative variable is added

$$\begin{aligned} \text{minimize} \quad & \boldsymbol{c}^\intercal \boldsymbol{x} + c_{n+1}x_{n+1} \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} + \boldsymbol{A}_{n+1}x_{n+1} = \boldsymbol{b}, \\ & \boldsymbol{x} \geq \boldsymbol{0}, \\ & x_{n+1} \geq 0. \end{aligned}$$

Then

$$\begin{bmatrix} \boldsymbol{x} \\ x_{n+1} \end{bmatrix} = \begin{bmatrix} \boldsymbol{x} \\ 0 \end{bmatrix}$$

is a basic feasible solution with the same basis matrix $\boldsymbol{B}$. This will be an optimal solution if and only if the additional optimality condition

$$\bar{c}_{n+1} = c_{n+1} - \boldsymbol{c}_B^\intercal \boldsymbol{B}^{-1}\boldsymbol{A}_{n+1} \geq 0$$

holds. If not, we add this variable and its reduced cost to the final simplex tableau for the original problem and continue simplex iterations from there.

Problem 5.1: Use the simplex method to solve the linear programming problem

$$\begin{array}{lrcrcrcrl} \text{minimize} & -5x_1 & - & x_2 & + & 12x_3 & & & \\ \text{subject to} & 3x_1 & + & 2x_2 & + & x_3 & & & = 10, \\ & 5x_1 & + & 3x_2 & & & + & x_4 & = 16, \\ & \multicolumn{8}{l}{x_1, x_2, x_3, x_4 \geq 0.} \end{array}$$

Two iterations of the simplex method (Phase II) lead us to the optimal solution $(2, 2, 0, 0)$, with optimal cost -12:

		x_1^*	x_2	x_3	x_4
	-120	-41	-25	0	0
$x_3 =$	10	3	2	1	0
$x_4^\dagger =$	16	5	3	0	1

		x_1	x_2^*	x_3	x_4
	$56/5$	0	$-2/5$	0	$41/5$
$x_3^\dagger =$	$2/5$	0	$1/5$	1	$-3/5$
$x_1 =$	$16/5$	1	$3/5$	0	$1/5$

		x_1	x_2	x_3	x_4
	12	0	0	2	7
$x_2 =$	2	0	1	5	-3
$x_1 =$	2	1	0	-3	2

Problem 5.2: Use an optimal solution of Problem 5.1 to construct an initial simplex tableau for the extended problem

$$\begin{array}{lrcrcrcrcrl} \text{minimize} & -5x_1 & - & x_2 & + & 12x_3 & & & - & x_5 & \\ \text{subject to} & 3x_1 & + & 2x_2 & + & x_3 & & & + & x_5 & = 10, \\ & 5x_1 & + & 3x_2 & & & + & x_4 & + & x_5 & = 16, \\ & \multicolumn{10}{l}{x_1, x_2, x_3, x_4, x_5 \geq 0,} \end{array}$$

in which a new variable x_5 has been added.

Recalling that the simplex tableau records $\boldsymbol{B}^{-1}\boldsymbol{A}$ at each iteration, and noting the final two columns of $\boldsymbol{A}$ form the identity matrix, we can simply read off

$$\boldsymbol{B}^{-1} = \begin{bmatrix} 5 & -3 \\ -3 & 2 \end{bmatrix}$$

from the final two columns of our tableau. Let us now introduce the extra variable x_5 into our tableau, entering

$$\boldsymbol{B}^{-1}\boldsymbol{A}_5 = \begin{bmatrix} 5 & -3 \\ -3 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

in the final column. We also need the reduced cost

$$\bar{c}_5 = c_5 - \boldsymbol{c}_B^\intercal \boldsymbol{B}^{-1}\boldsymbol{A}_5 = -1 - \begin{bmatrix} -1 & -5 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = -4.$$

Since $\bar{c}_5$ is negative, we see that $(2, 2, 0, 0, 0)$ is not an optimal solution to the extended problem. We therefore perform an additional simplex iteration:

			x_1	x_2	x_3	x_4	x_5^*
		12	0	0	2	7	-4
$x_2^\dagger$	=	2	0	1	5	-3	2
x_1	=	2	1	0	-3	2	-1

			x_1	x_2	x_3	x_4	x_5
		16	0	2	12	1	0
x_5	=	1	0	1/2	5/2	$-3/2$	1
x_1	=	3	1	1/2	$-1/2$	1/2	0

to obtain the optimal solution $(3, 0, 0, 0, 1)$ with optimal cost -16.

5.B. A new inequality constraint is added

$$\begin{aligned} \text{minimize} \quad & \boldsymbol{c}^\intercal \boldsymbol{x} \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \\ & \boldsymbol{a}_{m+1}^\intercal \boldsymbol{x} \geq b_{m+1}, \\ & \boldsymbol{x} \geq \boldsymbol{0}. \end{aligned}$$

If an optimal solution $\boldsymbol{x}^*$ to the original problem satisfies the inequality constraint, it is also an optimal solution to the new problem.

Otherwise, given an optimal basic feasible solution $\boldsymbol{x}^*$ to the original problem, we have $\boldsymbol{a}_{m+1}^\intercal \boldsymbol{x}^* < b_{m+1}$. We need to analyze the new problem by introducing a surplus variable x_{n+1}:

$$\begin{aligned} \text{minimize} \quad & \boldsymbol{c}^\intercal \boldsymbol{x} \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \\ & \boldsymbol{a}_{m+1}^\intercal \boldsymbol{x} - x_{n+1} = b_{m+1}, \\ & \boldsymbol{x} \geq \boldsymbol{0}, \\ & x_{n+1} \geq 0. \end{aligned}$$

Let $\bar{\boldsymbol{x}} = (\boldsymbol{x}, x_{n+1})$ and $\bar{\boldsymbol{b}} = (\boldsymbol{b}, b_{m+1})$. Then $\bar{\boldsymbol{A}}\bar{\boldsymbol{x}} = \bar{\boldsymbol{b}}$, where

$$\bar{\boldsymbol{A}} = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{0} \\ \boldsymbol{a}_{m+1}^\intercal & -1 \end{bmatrix}.$$

Let $\boldsymbol{B}$ be a basis matrix associated with $\boldsymbol{x}^*$, with basic components $\boldsymbol{x}_B^* = \boldsymbol{B}^{-1}\boldsymbol{b}$. Consider the basis matrix

$$\bar{\boldsymbol{B}} = \begin{bmatrix} \boldsymbol{B} & \boldsymbol{0} \\ \boldsymbol{a}_B^\intercal & -1 \end{bmatrix},$$

where $\boldsymbol{a}_B$ contains those elements of $\boldsymbol{a}_{m+1}$ associated with the original basic columns. It is straightforward to verify that the basic components of a basic solution to the new problem are given by

$$\begin{bmatrix} \boldsymbol{x}_B^* \\ \boldsymbol{a}_{m+1}^\intercal \boldsymbol{x}^* - b_{m+1} \end{bmatrix}.$$

This solution is infeasible since $\boldsymbol{a}_{m+1}^\intercal \boldsymbol{x}^* < b_{m+1}$. However, since

$$\bar{\boldsymbol{B}}^{-1} = \begin{bmatrix} \boldsymbol{B}^{-1} & \boldsymbol{0} \\ \boldsymbol{a}_B^\intercal \boldsymbol{B}^{-1} & -1, \end{bmatrix}$$

we can precompute

$$\bar{\boldsymbol{B}}^{-1}\bar{\boldsymbol{A}} = \begin{bmatrix} \boldsymbol{B}^{-1} & \boldsymbol{0} \\ \boldsymbol{a}_B^\intercal \boldsymbol{B}^{-1} & -1 \end{bmatrix} \begin{bmatrix} \boldsymbol{A} & \boldsymbol{0} \\ \boldsymbol{a}_{m+1}^\intercal & -1 \end{bmatrix} = \begin{bmatrix} \boldsymbol{B}^{-1}\boldsymbol{A} & \boldsymbol{0} \\ \boldsymbol{a}_B^\intercal \boldsymbol{B}^{-1}\boldsymbol{A} - \boldsymbol{a}_{m+1}^\intercal & 1 \end{bmatrix}$$

to find the new reduced cost vector

$$[\, \boldsymbol{c}^\intercal \quad 0\,] - [\, \boldsymbol{c}_B^\intercal \quad 0\,]\bar{\boldsymbol{B}}^{-1}\bar{\boldsymbol{A}} = [\, \boldsymbol{c}^\intercal - c_B^\intercal \boldsymbol{B}^{-1}\boldsymbol{A} \quad 0\,] \geq \boldsymbol{0}^\intercal.$$

As we saw in the proof of Theorem 4.4, the condition $\bar{\boldsymbol{c}} \geq \boldsymbol{0}$ for optimality of the primal solution is equivalent to the condition $\boldsymbol{A}^\intercal \boldsymbol{p} \leq \boldsymbol{c}$ for feasibility of the dual problem. This means that $\bar{\boldsymbol{B}}$ is a feasible basis for the dual problem. We can apply the so-called *dual simplex method*, described below, using the above values of the reduced costs and $\bar{\boldsymbol{B}}^{-1}\bar{\boldsymbol{A}}$ to construct the initial simplex tableau.

5.C. The dual simplex method

The *dual simplex method* can be used when one is given an initial tableau that has an infeasible basic solution, with basis matrix $\boldsymbol{B}$, but all reduced costs are nonnegative. Here is a single iteration of the dual simplex algorithm:

1. Examine the components of the vector $\boldsymbol{B}^{-1}\boldsymbol{b}$ in the zeroth column of the tableau. If they are all nonnegative, we have an optimal basic feasible solution and the algorithm terminates; otherwise, choose the lowest subscript j_ℓ such that $x_{j_\ell} < 0$.

2. The ℓth row of the tableau is the pivot row, with elements $x_{j_\ell}, v_1 \ldots, v_n$. If $v_j \geq 0$ for all j, the optimal cost of the dual problem is ∞ (so that the primal problem is infeasible) and the algorithm terminates.

3. Let j be the lowest subscript that minimizes $\{\bar{c}_j / {-v_j} : v_j < 0\}$. Column $\boldsymbol{A}_j$ will enter the basis, replacing column $\boldsymbol{A}_{j_\ell}$.

4. Add to each row of the tableau a multiple of the ℓth row (the pivot row) so that v_j (the pivot element) becomes 1 and all other entries of the pivot column become 0.

For example, the basic solution $(0, 0, 0, 2, -1)$ given by the following tableau is infeasible:

			x_1	x_2^*	x_3	x_4	x_5
		0	2	6	10	0	0
x_4	$=$	2	-2	4	1	1	0
$x_5^\dagger$	$=$	-1	4	-2	-3	0	1

However, since the reduced costs in the zeroth row are all nonnegative, we can apply the dual simplex method to obtain a basic feasible solution of the primal problem. After one iteration, we obtain the optimal solution $(0, 1/2, 0, 0, 0)$, with optimal cost 3:

			x_1	x_2	x_3	x_4	x_5
		-3	14	0	1	0	3
x_4	$=$	0	6	0	-5	1	2
x_2	$=$	1/2	-2	1	3/2	0	$-1/2$

Problem 5.3: Solve Prob. 5.1 with the added constraint $x_1 + x_2 \geq 5$.

With the added constraint, the linear programming problem can be written in standard form as

$$\begin{array}{llllllll} \text{minimize} & -5x_1 & - & x_2 & + & 12x_3 & & \\ \text{subject to} & 3x_1 & + & 2x_2 & + & x_3 & & = 10, \\ & 5x_1 & + & 3x_2 & & & + \; x_4 & = 16, \\ & x_1 & + & x_2 & & & - \; x_5 & = 5, \\ & \multicolumn{7}{l}{x_1, x_2, x_3, x_4, x_5 \geq 0.} \end{array}$$

Let $a_{m+1} = (1, 1, 0, 0)$ and $b_{m+1} = 5$. In the final tableau of the original problem, x_2 and x_1 are the basic variables, so $\boldsymbol{B} = \{\boldsymbol{A}_2 \; \boldsymbol{A}_1\}$. Thus $\boldsymbol{a}_B = (1, 1)$. We first compute

$$\boldsymbol{a}_B^\intercal \boldsymbol{B}^{-1} \boldsymbol{A} - \boldsymbol{a}_{m+1}^\intercal = [1 \quad 1] \begin{bmatrix} 0 & 1 & 5 & -3 \\ 1 & 0 & -3 & 2 \end{bmatrix} - [1 \quad 1 \quad 0 \quad 0] = [0 \quad 0 \quad 2 \quad -1].$$

For the given basic solution $(2, 2, 0, 0)$ of the original problem, we see that the new slack variable x_5 must be -1 in order to satisfy the added constraint. Since $x_5 < 0$, the basic solution to the extended problem is infeasible. We then need to apply the dual simplex method to the tableau

			x_1	x_2	x_3	x_4^*	x_5
		12	0	0	2	7	0
x_2	$=$	2	0	1	5	-3	0
x_1	$=$	2	1	0	-3	2	0
$x_5^\dagger$	$=$	-1	0	0	2	-1	1

Since v_4 is the only negative value in the row for x_5, the corresponding variable x_4 enters and x_5 exits. One iteration of the dual simplex method then brings us to the optimal solution $(0,5,0,1,0)$ with optimal cost -5:

			x_1	x_2	x_3	x_4	x_5
		5	0	0	16	0	7
x_2	$=$	5	0	1	-1	0	-3
x_1	$=$	0	1	0	1	0	2
x_4	$=$	1	0	0	-2	1	-1

5.D. Changes in the target vector $\boldsymbol{b}$

Assume that $\boldsymbol{b}$ is changed to $\boldsymbol{b}+\delta\boldsymbol{e}_k$ for some $k \in \{1,\ldots,m\}$. We want to determine the range of values of δ under which the current basis remains optimal. Since the optimality condition $\bar{\boldsymbol{c}} \geq \boldsymbol{0}$ is unchanged, we only need to check the feasibility condition

$$\bar{\boldsymbol{x}}_B = \boldsymbol{B}^{-1}(\boldsymbol{b}+\delta\boldsymbol{e}_k) = \boldsymbol{x}_B + \delta\boldsymbol{B}^{-1}\boldsymbol{e}_k \geq \boldsymbol{0}.$$

Let $\boldsymbol{g} = \boldsymbol{B}^{-1}\boldsymbol{e}_k = (\beta_1,\ldots,\beta_m)$ be the kth column of $\boldsymbol{B}^{-1}$. Feasibility will be maintained if

$$\boldsymbol{x}_B + \delta\boldsymbol{g} \geq \boldsymbol{0},$$

that is, if

$$\max_{i:\beta_i>0}\left(-\frac{x_{j_i}}{\beta_i}\right) \leq \delta \leq \min_{i:\beta_i<0}\left(-\frac{x_{j_i}}{\beta_i}\right).$$

The optimal cost is then given by

$$\boldsymbol{c}_B^{\mathsf{T}}\bar{\boldsymbol{x}}_B = \boldsymbol{c}_B^{\mathsf{T}}\boldsymbol{B}^{-1}(\boldsymbol{b}+\delta\boldsymbol{e}_k) = \boldsymbol{p}^{\mathsf{T}}\boldsymbol{b} + \delta p_k,$$

where $\boldsymbol{p}^{\mathsf{T}} = \boldsymbol{c}_B^{\mathsf{T}}\boldsymbol{B}^{-1}$ is an optimal solution of the dual problem. If δ is outside the above range, the current solution satisfies the optimality (or dual feasibility) conditions but is prime infeasible. Then one can apply the dual simplex algorithm, starting with the current basis.

Problem 5.4: Consider the optimal tableau in Problem 5.1. If we add δ to b_1, for which range of δ does the solution remain optimal? Within this range, what is the rate of change of the optimal cost per unit δ?

We only need check that our basic solution remains feasible:

$$\bar{\boldsymbol{x}}_B = \boldsymbol{B}^{-1}(\boldsymbol{b}+\delta\boldsymbol{e}_1) = \begin{bmatrix} 5 & -3 \\ -3 & 2 \end{bmatrix}\begin{bmatrix} 10+\delta \\ 16 \end{bmatrix} = \begin{bmatrix} 2+5\delta \\ 2-3\delta \end{bmatrix}.$$

For both components of $\bar{\boldsymbol{x}}_B$ to remain nonnegative, we require $\delta \in [-2/5, 2/3]$. The rate of change of the optimal cost per unit δ is

$$\boldsymbol{c}_B^{\mathsf{T}}\boldsymbol{B}^{-1}\boldsymbol{e}_1 = [\,-1 \quad -5\,]\begin{bmatrix} 5 \\ -3 \end{bmatrix} = 10.$$

For example, if we choose $\delta = -2/5$, our total cost would change by $-2/5 \times 10 = -4$ (a cost reduction of 4).

5.E. Changes in the cost vector $\boldsymbol{c}$

Assume that c_j changes to $c_j + \delta$ for some index j. Prime feasibility is not affected. We therefore focus on the optimality condition $\boldsymbol{c}^\intercal - \boldsymbol{c}_B^\intercal \boldsymbol{B}^{-1}\boldsymbol{A} \geq \boldsymbol{0}^\intercal$.

Case 1: If x_j is a nonbasic variable, $\boldsymbol{c}_B$ does not change. The only inequality that is affected is the one for the reduced cost of moving in the jth simplex direction:

$$c_j + \delta - \boldsymbol{c}_B^\intercal \boldsymbol{B}^{-1}\boldsymbol{A}_j \geq 0.$$

Thus, if $\delta \geq \boldsymbol{c}_B^\intercal \boldsymbol{B}^{-1}\boldsymbol{A}_j - c_j = -\bar{c}_j$, the current basis remains optimal; otherwise, we can apply the primal simplex method starting from the current basic feasible solution.

Case 2: If x_j is the ℓth basic variable, i.e. $j = j_\ell$, then $\boldsymbol{c}_B$ becomes $\boldsymbol{c}_B + \delta \boldsymbol{e}_\ell$. The optimality conditions for the new problem are then

$$0 \leq c_i - (\boldsymbol{c}_B + \delta\boldsymbol{e}_\ell)^\intercal \boldsymbol{B}^{-1}\boldsymbol{A}_i = \bar{c}_i - \delta q_i \quad \text{for all } i \neq j,$$

where $\boldsymbol{q}$ is the ℓth row of $\boldsymbol{B}^{-1}\boldsymbol{A}$ from the simplex tableau. Since $\boldsymbol{B}^{-1}\boldsymbol{B} = \boldsymbol{1}$, the tableau entry q_i corresponding to a basic variable x_i is δ_{ij}. When $i \neq j$ the new reduced cost of the basic variable x_i is $\bar{c}_i - \delta q_i = 0 - \delta\delta_{ij} = 0$. When $i = j$, the new reduced cost of the basic variable x_j is $(\bar{c}_j + \delta) - \delta q_j = 0 + \delta - \delta\delta_{jj} = \delta - \delta = 0$. That is, the reduced cost of each of the basic variables remains zero. Preserving optimality therefore only requires that

$$q_i \delta \leq \bar{c}_i \quad \text{for all nonbasic indices } i.$$

Problem 5.5: For $j = 1, \ldots 4$, determine the range of change δ_j of c_j under which the final basis in Problem 5.1 remains optimal.

We note that $\bar{\boldsymbol{c}} = (0, 0, 2, 7)$, with the nonbasic variables being x_3 and x_4. So we require

$$\delta_3 \geq -2;$$

$$\delta_4 \geq -7.$$

For the basic variable x_1, we see that $\boldsymbol{q} = (1, 0, -3, 2)$. We require that $q_i \delta_1 \leq \bar{c}_i$ for the nonbasic indices $i = 3$ and $i = 4$:

$$-3\delta_1 \leq 2;$$

$$2\delta_1 \leq 7.$$

The solution will thus remain optimal as long as $\delta_1 \in [-2/3, 7/2]$.

For the basic variable x_2, we see that $\boldsymbol{q} = (0, 1, 5, -3)$. We require that $q_i \delta_2 \leq \bar{c}_i$ for the nonbasic indices $i = 3$ and $i = 4$:

$$5\delta_2 \leq 2;$$

$$-3\delta_2 \leq 7.$$

The solution will thus remain optimal as long as $\delta_2 \in [-7/3, 2/5]$.

Chapter 6

Parametric Programming

Sometimes a linear programming problem contains unknown parameters. The simplex method can nevertheless still be used, if algebraic (symbolic) computations are substituted for numerical calculations. Various cases corresponding to different parameter regimes may arise. For example, consider the linear programming problem

$$\begin{array}{llllllll} \text{minimize} & (-3+2t)x_1 & + & (3-t)x_2 & + & x_3 & & \\ \text{subject to} & x_1 & + & 2x_2 & - & 3x_3 & \leq 5, \\ & 2x_1 & + & x_2 & - & 4x_3 & \leq 7, \\ & x_1, x_2, x_3 \geq 0. \end{array}$$

We introduce slack variables x_4 and x_5 to put the problem into standard form:

$$\begin{array}{llllllllll} \text{minimize} & (-3+2t)x_1 & + & (3-t)x_2 & + & x_3 & & & & \\ \text{subject to} & x_1 & + & 2x_2 & - & 3x_3 & + & x_4 & & = 5, \\ & 2x_1 & + & x_2 & - & 4x_3 & & & + \; x_5 & = 7, \\ & x_1, x_2, x_3, x_4, x_5 \geq 0. \end{array}$$

An initial simplex tableau is then

			x_1	x_2	x_3	x_4	x_5
		0	$-3+2t$	$3-t$	1	0	0
x_4	=	5	1	2	-3	1	0
x_5	=	7	2	1	-4	0	1

Our analysis then bifurcates into various cases.

Case 1: $t \in [3/2, 3]$: Since $-3+2t \geq 0$ and $3-t \geq 0$, the simplex method terminates with the optimal solution $\boldsymbol{x}^* = (0,0,0,5,7)$ and optimal cost 0.

Case 2: $t > 3$. We see that x_2 enters and x_4 exits, yielding

			x_1	x_2	x_3	x_4	x_5
		$(5t-15)/2$	$(5t-9)/2$	0	$(11-3t)/2$	$(t-3)/2$	0
x_2	=	$5/2$	$1/2$	1	$-3/2$	$1/2$	0
x_5	=	$9/2$	$3/2$	0	$-5/2$	$-1/2$	1

Case 2.1: $t \in (3, 11/3]$: The simplex method terminates with the optimal solution $\boldsymbol{x}^* = (0, 5/2, 0, 0, 9/2)$ and optimal cost $(15 - 5t)/2$.

Case 2.2: $t > 11/3$: The problem is unbounded, with optimal cost $-\infty$.

Case 3: $t < 3/2$. Then x_1 enters and x_5 exits, yielding

		x_1	x_2	x_3	x_4	x_5
	$21/2 - 7t$	0	$9/2 - 2t$	$-5 + 4t$	0	$3/2 - t$
$x_4 =$	$3/2$	0	$3/2$	-1	1	$-1/2$
$x_1 =$	$7/2$	1	$1/2$	-2	0	$1/2$

Case 3.1: $t \in [5/4, 3/2]$: The simplex method terminates with the optimal solution $\boldsymbol{x}^* = (7/2, 0, 0, 3/2, 0)$ and optimal cost is $7t - 21/2$.

Case 3.2: $t < 5/4$: The problem is unbounded, with optimal cost $-\infty$.

Bibliography

[Bertsimas & Tsitsiklis 1997] D. Bertsimas & J. N. Tsitsiklis, *Introduction to linear optimization*, volume 6, Athena Scientific, 1997.

Index