MA427 – Mathematical Optimisation
Part I
Linear Programming

Giacomo Zambelli
g.zambelli@lse.ac.uk

Department of Mathematics
London School of Economics and Political Science
2018/19

Contents

Chapter 1

Systems of Linear Inequalities

We start by studying systems of linear inequalities $Ax \le b$. We look at this subject from two different angles. In this chapter, we use a more algebraic approach to address the issue of solvability of $Ax \le b$. In the next chapter, we take a geometric viewpoint to study the geometric properties of the set of solutions $\{x \in \mathbb{R}^n \mid Ax \le b\}$ of such systems.

Notation and basic linear algebra

By default, by vectors will mean column vectors. For an m dimensional vector $v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{pmatrix}$, its transpose $v^\top = (v_1, v_2, \dots, v_m)$ is a row vector.

Throughout the notes, for a matrix $A \in \mathbb{R}^{m \times n}$, we will denote by A_j the jth column of A, $j = 1, \dots, n$ (therefore A_j is a vector in $\mathbb{R}^m$).

The rows of A will be denoted by $a_1^\top, \dots, a_m^\top$ (thus a_i is a vector in $\mathbb{R}^n$ for $i = 1, \dots, m$).

Recall from linear algebra that the *rank* of a matrix A, which we denote $\mathrm{rk}(A)$, is the largest number of rows of A that are linearly independent. We recall also the basic result that the largest number of rows of A that are linearly independent is equal to the largest number of columns of A that are linearly independent (that is $\mathrm{rk}(A) = \mathrm{rk}(A^\top)$).

A system of linear equations is of the form

$$Ax = b,$$

where $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. We can use the *Gaussian elimination* method to find a feasible solution x, or to conclude that the system is infeasible.

We recall that $Ax = b$ is feasible if and only if $\mathrm{rk}(A) = \mathrm{rk}(A|b)$, where the matrix $(A|b) \in \mathbb{R}^{m \times (n+1)}$ is obtained by juxtaposing the column vector b to the matrix A.

1.1 Fourier elimination

The most basic question concerning a system of linear inequalities is whether or not it has a solution. Fourier devised a simple method to address this problem. Fourier's method is similar to Gaussian elimination, in that it performs row operations to eliminate one variable at a time.

Let $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$, and suppose we want to determine if the system $Ax \le b$ has a solution. We first reduce this question to one about a system with $n - 1$ variables. Namely, we determine necessary and sufficient conditions for which, given a vector $(\bar{x}_1, \dots, \bar{x}_{n-1}) \in \mathbb{R}^{n-1}$, there exists $\bar{x}_n \in \mathbb{R}$ such that $(\bar{x}_1, \dots, \bar{x}_n)$ satisfies $Ax \le b$. Let $I := \{1, \dots, m\}$ and define

$$I^+ := \{i \in I \mid a_{in} > 0\}, \quad I^- := \{i \in I \mid a_{in} < 0\}, \quad I^0 := \{i \in I \mid a_{in} = 0\}.$$

Dividing the ith row by $|a_{in}|$ for each $i \in I^+ \cup I^-$, we obtain the following system, which is equivalent to $Ax \leq b$:

$$\begin{array}{llll} \sum_{j=1}^{n-1} a'_{ij}x_j & +x_n & \leq b'_i, & i \in I^+ \\ \sum_{j=1}^{n-1} a'_{ij}x_j & -x_n & \leq b'_i, & i \in I^- \\ \sum_{j=1}^{n-1} a_{ij}x_j & & \leq b_i, & i \in I^0 \end{array} \tag{1.1}$$

where $a'_{ij} = a_{ij}/|a_{in}|$ and $b'_i = b_i/|a_{in}|$ for $i \in I^+ \cup I^-$.

For each pair $i \in I^+$ and $k \in I^-$, we sum the two inequalities indexed by i and k, and we add the resulting inequality to the system (1.1). Furthermore, we remove the inequalities indexed by I^+ and I^-. This way, we obtain the following system:

$$\begin{array}{rcll} \sum_{j=1}^{n-1}(a'_{ij} + a'_{kj})x_j & \leq & b'_i + b'_k, & i \in I^+,\ k \in I^-, \\ \sum_{j=1}^{n-1} a_{ij}x_j & \leq & b_i, & i \in I^0. \end{array} \tag{1.2}$$

If $(\bar{x}_1, \ldots, \bar{x}_{n-1}, \bar{x}_n)$ satisfies $Ax \leq b$, then $(\bar{x}_1, \ldots, \bar{x}_{n-1})$ satisfies (1.2). The next theorem states that the converse also holds.

Theorem 1.1. *A vector $(\bar{x}_1, \ldots, \bar{x}_{n-1})$ satisfies the system (1.2) if and only if there exists $\bar{x}_n$ such that $(\bar{x}_1, \ldots, \bar{x}_{n-1}, \bar{x}_n)$ satisfies $Ax \leq b$.*

Proof. We already remarked the "if" statement. For the converse, assume there is a vector $(\bar{x}_1, \ldots, \bar{x}_{n-1})$ satisfying (1.2). Note that the first set of inequalities in (1.2) can be rewritten as

$$\sum_{j=1}^{n-1} a'_{kj}x_j - b'_k \leq b'_i - \sum_{j=1}^{n-1} a'_{ij}x_j, \quad i \in I^+,\ k \in I^-. \tag{1.3}$$

Let $l := \max_{k \in I^-}\{\sum_{j=1}^{n-1} a'_{kj}\bar{x}_j - b'_k\}$ and $u := \min_{i \in I^+}\{b'_i - \sum_{j=1}^{n-1} a'_{ij}\bar{x}_j\}$, where we define $l := -\infty$ if $I^- = \emptyset$ and $u := +\infty$ if $I^+ = \emptyset$. Since $(\bar{x}_1, \ldots, \bar{x}_{n-1})$ satisfies (1.3), we have that $l \leq u$. Therefore, for any $\bar{x}_n$ such that $l \leq \bar{x}_n \leq u$, the vector $(\bar{x}_1, \ldots, \bar{x}_n)$ satisfies the system (1.1), which is equivalent to $Ax \leq b$. □

Therefore, the problem of finding a solution to $Ax \leq b$ is reduced to finding a solution to (1.2), which is a system of linear inequalities in $n-1$ variables. *Fourier's elimination method* is:

Given a system of linear inequalities $Ax \leq b$, let $A^n := A$, $b^n := b$;
For $i = n, \ldots, 1$, eliminate variable x_i from $A^i x \leq b^i$ with the above procedure to obtain system $A^{i-1}x \leq b^{i-1}$.

System $A^1x \leq b^1$, which involves variable x_1 only, is of the type, $x_1 \leq b^1_p$, $p \in P$, $-x_1 \leq b^1_q$, $q \in N$, and $0 \leq b^1_i$, $i \in Z$.

System $A^0x \leq b^0$ has the following inequalities: $0 \leq b^0_{pq} := b^1_p + b^1_q$, $p \in P$, $q \in N$, $0 \leq b^0_i := b^1_i$, $i \in Z$.

Applying Theorem 1.1, we obtain that $Ax \leq b$ is feasible if and only if $A^0x \leq b^0$ is feasible, and this happens when the b^0_{pq} and b^0_i are all nonnegative.

Remark 1.2. At each iteration, Fourier's method removes $|I^+| + |I^-|$ inequalities and adds $|I^+| \times |I^-|$ inequalities, hence the number of inequalities may roughly be squared at each iteration. Thus, after eliminating p variables, the number of inequalities may be exponential in p. Therefore, despite its theoretical importance and simplicity, Fourier's method is not commonly used for solving systems of linear inequalities.

Example 1.3. Consider the system $A^3x \leq b^3$ of linear inequalities in three variables

$$\begin{array}{rrrrrrrr} - & x_1 & & & & & \leq & -1 \\ & & - & x_2 & & & \leq & -1 \\ & & & & - & x_3 & \leq & -1 \\ - & x_1 & - & x_2 & & & \leq & -3 \\ - & x_1 & & & - & x_3 & \leq & -3 \\ & & - & x_2 & - & x_3 & \leq & -3 \\ & x_1 & + & x_2 & + & x_3 & \leq & 6 \end{array}$$

Applying Fourier's procedure to eliminate variable x_3, we obtain the system $A^2x \le b^2$:

$$\begin{array}{rrrrl}
- & x_1 & & & \le & -1 \\
 & & - & x_2 & \le & -1 \\
- & x_1 & - & x_2 & \le & -3 \\
 & x_1 & + & x_2 & \le & 5 \\
 & & & x_2 & \le & 3 \\
 & x_1 & & & \le & 3
\end{array}$$

where the last three inequalities are obtained from $A^3x \le b^3$ by summing the third, fifth, and sixth inequality, respectively, with the last inequality. Eliminating variable x_2, we obtain $A^1x \le b^1$

$$\begin{array}{rrll}
- & x_1 & \le & -1 \\
 & x_1 & \le & 3 \\
 & x_1 & \le & 4 \\
 & 0 & \le & 2 \\
 & 0 & \le & 2 \\
- & x_1 & \le & 0
\end{array}$$

Finally $A^0x \le b^0$ is

$$\begin{array}{rlr}
0 & \le & 3-1 \\
0 & \le & 4-1 \\
0 & \le & 3 \\
0 & \le & 4 \\
0 & \le & 2 \\
0 & \le & 2
\end{array}$$

Therefore $A^0x \le b^0$ is feasible. A solution can now be found by *backward substitution*. System $A^1x \le b^1$ is equivalent to $1 \le x_1 \le 3$. Since x_1 can take any value in this interval, choose $\bar{x}_1 = 3$. Substituting $x_1 = 3$ in $A^2x \le b^2$, we obtain $1 \le x_2 \le 2$. If we choose $\bar{x}_2 = 1$ and substitute $x_2 = 1$ and $x_1 = 3$ in $A^3x \le b^3$, we finally obtain $x_3 = 2$. This gives the solution $\bar{x} = (3, 1, 2)$.

Expressing the derived constraints as nonnegative combinations

The following simple observation will be very useful.

Proposition 1.4. *Every inequality of $A^ix \le b^i$ is a nonnegative combination of inequalities of $Ax \le b$. That is, for every inequality $c^\top x \le d$ in the system $A^ix \le b^i$, we can find a nonnegative vector $u \in \mathbb{R}^m$, $u \ge 0$, such that $c = u^\top A$, and $d = u^\top b$.*

The coefficients in u can be determined by tracing back how $c^\top x \le d$ was obtained during the Fourier elimination process. Let us illustrate this on the previous example. We start with $A = A^3$ and $b = b^3$. Every inequality in the system $A^2x \le b^2$ is either an original inequality from $Ax \le b$, or is obtained by adding two inequalities from $Ax \le b$. For example, the second inequality $-x_2 \le -1$ of $A^2x \le b^2$ is already present in $Ax \le b$, whereas $x_1 + x_2 \le 5$ is obtained from adding the inequalities $-x_3 \le -1$ and $x_1 + x_2 + x_3 \le 6$.

The inequality $-x_2 \le -1$ is of the form $c^\top x \le d$ for $c^\top = (0, -1, 0)$ and $d = -1$. This can be obtained as $c = u^\top A$ and $d = u^\top b$ for $u^\top = (0, 1, 0, 0, 0, 0, 0)$.

Similarly, the inequality $x_1 + x_2 \le 5$ is of the form $c^\top x \le d$ for $c^\top = (1, 1, 0)$ and $d = 5$. We obtain this as $c = u^\top A$ and $d = u^\top b$ for $u^\top = (0, 0, 1, 0, 0, 0, 1)$.

Let us now move to the next system $A^1x \le b^1$. The third inequality, $x_1 \le 4$ was obtained by adding $-x_2 \le -1$ and $x_1 + x_2 \le 5$. We can write it as $c^\top x \le d$ for $c^\top = (1, 0, 0)$ and $d = 4$. We obtain this in the form $c = u^\top A$ and $d = u^\top b$ for $u = (0, 1, 1, 0, 0, 0, 1)$.

1.2 Farkas' Lemma

Next we present Farkas' lemma, which gives a simple necessary and sufficient condition for the existence of a solution to a system of linear inequalities. The proof is based on Fourier elimination.

Theorem 1.5 (Farkas' lemma)**.** *A system of linear inequalities $Ax \leq b$ is infeasible if and only if the system $u^\top A = 0,\ u^\top b < 0,\ u \geq 0$ is feasible.*

Proof. Assume $u^\top A = 0,\ u^\top b < 0,\ u \geq 0$ is feasible. Then $0 = u^\top Ax \leq u^\top b < 0$ for any x satisfying $Ax \leq b$, since it is a nonnegative combination of inequalities in the system. It follows that $Ax \leq b$ is infeasible and this proves the "if" part.

We now prove the "only if" part. Assume that $Ax \leq b$ has no solution. Apply the Fourier elimination method to $Ax \leq b$ to eliminate all variables $x_n, \dots, x_1$. System $A^0x \leq b^0$ is of the form $0 \leq b^0$, and the system $Ax \leq b$ has a solution if and only if all the entries of b^0 are nonnegative. Since $Ax \leq b$ has no solution, it follows that b^0 has a negative entry, say $b^0_i < 0$. By Proposition 1.4, every inequality of the system $0 \leq b^0$ is a nonnegative combination of inequalities of $Ax \leq b$. In particular, there exists some vector $u \geq 0$ such that the inequality $0 \leq b^0_i$ is identical to $u^\top Ax \leq u^\top b$. That is, $u \geq 0$, $u^\top A = 0$, $u^\top b = b^0_i < 0$ is feasible. □

Farkas' lemma is sometimes referred to as a *theorem of the alternative* because it can be restated as follows.

Exactly one among the system $Ax \leq b$ and the system $u^\top A = 0,\ u^\top b < 0,\ u \geq 0$ is feasible.

The following is Farkas' lemma for systems of equations in nonnegative variables.

Theorem 1.6. *The system $Ax = b,\ x \geq 0$ is feasible if and only if the system $u^\top A \leq 0,\ u^\top b > 0$ is infeasible.*

Proof. Assume first that $Ax = b,\ x \geq 0$ is feasible, and consider a vector u such that $u^\top A \leq 0$. Then $u^\top b = u^\top(Ax) = (u^\top A)x \leq 0$, since $u^\top A$ is a nonpositive and x is a nonnegative vector. Consequently, the system $u^\top A \leq 0$, $u^\top b > 0$ is infeasible.

For the converse, suppose that $Ax = b,\ x \geq 0$ is infeasible. We can write this system equivalently as $Ax \leq b,\ -Ax \leq -b,\ -x \leq 0$. Since this is infeasible, by Theorem 1.5 there exists $(v, v', w) \geq 0$ such that $v^\top A - {v'}^\top A - w = 0$ and $v^\top b - {v'}^\top b < 0$. The vector $u := v' - v$ satisfies $u^\top b > 0$ and since $w \geq 0$, u satisfies $u^\top A \leq 0$. □

1.3 Linear programming duality

Linear programming is the problem of maximizing a linear function subject to a finite number of linear constraints. Given a matrix $A \in \mathbb{R}^{m \times n}$ and vectors $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, the *dual* of the linear programming problem $\max\{c^\top x \mid Ax \leq b\}$ is the problem $\min\{u^\top b \mid u^\top A = c,\ u \geq 0\}$. Next we derive the fundamental theorem of linear programming, stating that the optimum values of the primal and dual problems coincide whenever both problems have a feasible solution. This property is called *strong duality.*

Theorem 1.7 (Strong Duality Theorem)**.** *Given a matrix $A \in \mathbb{R}^{m \times n}$ and vectors $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, let $P := \{x \mid Ax \leq b\}$ and $D := \{u \mid u^\top A = c,\ u \geq 0\}$. If P and D are both nonempty, then*

$$\max\{c^\top x \mid Ax \leq b\} = \min\{u^\top b \mid u^\top A = c,\ u \geq 0\}, \tag{1.4}$$

and there exist $x^* \in P$ and $y^* \in D$ such that $c^\top x^* = {u^*}^\top b$.

Proof. For every $x \in P$ and $u \in D$, we have $c^\top x = u^\top Ax \leq u^\top b$, where the equality follows from $u^\top A = c$ and the inequality follows from $u \geq 0$, $Ax \leq b$. Hence $\max\{c^\top x \mid x \in P\} \leq \min\{u^\top b \mid u \in D\}$. Since $D \neq \emptyset$, this also implies that $\max\{c^\top x \mid x \in P\}$ is bounded.

Note that $\max\{c^\top x \mid x \in P\} = \max\{z \mid z - c^\top x \leq 0, Ax \leq b\}$. Apply Fourier's method to the system $z - c^\top x \leq 0$, $Ax \leq b$, to eliminate the variables $x_1, \ldots, x_n$. The result is a system $\bar{a}z \leq \bar{b}$ in the single variable z, where $\bar{a}$ and $\bar{b}$ are vectors. We may assume that the entries of $\bar{a}$ are $0, \pm 1$. By Theorem 1.1, $\max\{z \mid z - c^\top x \leq 0, Ax \leq b\} = \max\{z \mid \bar{a}z \leq \bar{b}\}$, and there exists $x^* \in P$ such that $c^\top x^*$ achieves the maximum. Since $\max\{z \mid \bar{a}z \leq \bar{b}\}$ is bounded, at least one entry of $\bar{a}$ equals 1, and $\max\{z \mid \bar{a}z \leq \bar{b}\} = \min_{i \mid \bar{a}_i = 1} \bar{b}_i$. Let h be an index achieving the minimum in the previous equation. By Proposition 1.4, inequality $z \leq \bar{b}_h$ is a nonnegative combination of the $m+1$ inequalities $z - c^\top x \leq 0$, $Ax \leq b$. Thus there exists a nonnegative vector $(u_0, u^*) \in \mathbb{R}_+ \times \mathbb{R}^m_+$ such that $u_0 = 1$, $(u_0, u^*)^\top \binom{-c}{A} = 0$ and ${u^*}^\top b = \bar{b}_h$. It follows that $u^* \in D$ and ${u^*}^\top b = \max\{c^\top x \mid x \in P\}$. □

Recall that $a_i^\top$ denotes the ith row of the matrix $A \in \mathbb{R}^{m \times n}$.

Theorem 1.8 (Complementary slackness). *Given a matrix $A \in \mathbb{R}^{m \times n}$ and vectors $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, let $P := \{x \mid Ax \leq b\}$ and $D := \{u \mid u^\top A = c,\ u \geq 0\}$. Given $x^* \in P$ and $u^* \in D$, x^* and u^* are optimal solutions for the primal and dual problem* $\max\{c^\top x \mid x \in P\}$ *and* $\min\{u^\top b \mid u \in D\}$*, respectively, if and only if the following* complementary slackness conditions *hold*

$$u_i^*({a_i}^\top x^* - b_i) = 0 \text{ for } i = 1, \ldots, m.$$

Proof. We have that $c^\top x^* = {u^*}^\top A x^* \leq {u^*}^\top b$, and by Theorem 1.7 equality holds if and only if x^* and u^* are optimal solutions for $\max\{c^\top x \mid x \in P\}$ and $\min\{u^\top b \mid u \in D\}$. Since ${a_i}^\top x^* \leq b_i$ and $u_i^* \geq 0$, equality holds if and only if, for $i = 1, \ldots, m$, $u_i^*(a_i^\top x^* - b_i) = 0$. □

Here is another consequence of Farkas' lemma:

Proposition 1.9. *Let $P := \{x \mid Ax \leq b\}$ and $D := \{u \mid u^\top A = c,\ u \geq 0\}$, and suppose $P \neq \emptyset$. Then* $\max\{c^\top x \mid x \in P\}$ *is unbounded if and only if $D = \emptyset$. Equivalently,* $\max\{c^\top x \mid x \in P\}$ *is unbounded if and only if there exists a vector $\bar{y}$ such that $A\bar{y} \leq 0$ and $c^\top \bar{y} > 0$.*

Proof. By Farkas' lemma, $D = \emptyset$ if and only if there exists a vector $\bar{y}$ such that $A\bar{y} \leq 0$ and $c^\top \bar{y} > 0$. If $D \neq \emptyset$, then by Theorem 1.7 $\max\{c^\top x \mid x \in P\} = \min\{u^\top b \mid u \in D\}$, therefore $\max\{c^\top x \mid x \in P\}$ is bounded. Conversely, assume $D = \emptyset$. By Theorem 1.6, there exists a y such that $Ay \leq 0$ and $c^\top y > 0$. Given $\bar{x} \in P$ and such a $\bar{y}$, it follows that $\bar{x} + \lambda\bar{y} \in P$ for every $\lambda \geq 0$ and $\lim_{\lambda \to +\infty} c^\top(\bar{x} + \lambda\bar{y}) = +\infty$. Thus $\max\{c^\top x \mid x \in P\}$ is unbounded. □

Chapter 2

Geometry of Linear Programming

2.1 Linear, convex, and conic combinations

A *hyperplane* in $\mathbb{R}^n$ is a set of the form $\{x \in \mathbb{R}^n \mid a^\top x = \beta\}$, where a is a nonzero vector of $\mathbb{R}^n$ and $\beta \in \mathbb{R}$.
A *half-space* in $\mathbb{R}^n$ is a set of the form $\{x \in \mathbb{R}^n \mid a^\top x \leq \beta\}$, where a is a nonzero vector of $\mathbb{R}^n$ and $\beta \in \mathbb{R}$.
A *polyhedron* in $\mathbb{R}^n$ is the intersection of a finite number of half-spaces. Equivalently, a polyhedron is a set that can be written in the form $P = \{x \in \mathbb{R}^n \mid A'x = b',\ A''x \leq b''\}$ where $A' \in \mathbb{R}^{k\times n}$, $b' \in \mathbb{R}^k$, $A'' \in \mathbb{R}^{(m-k)\times n}$, $b'' \in \mathbb{R}^{(m-k)}$. Therefore, the feasible region of a Linear Programming problem is a polyhedron.

Linear combinations, linear spaces

Vector $x \in \mathbb{R}^n$ is a *linear combination* of the vectors $x^1, \ldots, x^q \in \mathbb{R}^n$ if there exist scalars $\lambda_1, \ldots, \lambda_q$ such that

$$x = \sum_{j=1}^{q} \lambda_j x^j.$$

Vectors $x^1, \ldots, x^q \in \mathbb{R}^n$ are *linearly independent* if $\lambda_1 = \ldots = \lambda_q = 0$ is the unique solution to the system $\sum_{j=1}^{q} \lambda_j x^j = 0$.

A nonempty subset $\mathcal{L}$ of $\mathbb{R}^n$ is a *linear space* if $\mathcal{L}$ is closed under taking linear combinations, i.e. every linear combination of vectors in $\mathcal{L}$ belongs to $\mathcal{L}$. A subset $\mathcal{L}$ of $\mathbb{R}^n$ is a linear space if and only if $\mathcal{L} = \{x \in \mathbb{R}^n \mid Ax = 0\}$ for some matrix A.

A *basis* of a linear space $\mathcal{L}$ is a maximal set of linearly independent vectors in $\mathcal{L}$. All bases have the same cardinality. This cardinality is called the *dimension* of $\mathcal{L}$. If $\mathcal{L} = \{x \in \mathbb{R}^n \mid Ax = 0\}$, then the dimension of $\mathcal{L}$ is $n - \text{rank}(A)$. We note that a hyperplane in $\mathbb{R}^n$ is an $n-1$ dimensional linear subspace.

The inclusionwise minimal linear space containing a set $S \subseteq \mathbb{R}^n$ is the *linear space generated* by S, and is denoted by $\langle S \rangle$. Given any maximal set S' of linearly independent vectors in S, we have that $\langle S \rangle = \langle S' \rangle$.

Affine combinations, affine spaces

A point $x \in \mathbb{R}^n$ is an *affine combination* of $x^1, \ldots, x^q \in \mathbb{R}^n$ if there exist scalars $\lambda_1, \ldots, \lambda_q$ such that

$$x = \sum_{j=1}^{q} \lambda_j x^j, \quad \sum_{j=1}^{q} \lambda_j = 1.$$

Points $x^0, x^1, \ldots, x^q \in \mathbb{R}^n$ are *affinely independent* if $\lambda_0 = \lambda_1 = \ldots = \lambda_q = 0$ is the unique solution to the system

$$\sum_{j=0}^{q} \lambda_j x^j = 0, \quad \sum_{j=0}^{q} \lambda_j = 0.$$

Equivalently, $x^0, x^1, \ldots, x^q \in \mathbb{R}^n$ are affinely independent if and only if no point in $x^0, \ldots, x^q$ can be written as an affine combination of the others.

A subset $\mathcal{A}$ of $\mathbb{R}^n$ is an *affine space* if $\mathcal{A}$ is closed under taking affine combinations. A *basis of an affine subspace* $\mathcal{A} \subseteq \mathbb{R}^n$ is a maximal set of affinely independent points in $\mathcal{A}$. All bases of $\mathcal{A}$ have the same cardinality.

Equivalently, $\mathcal{A} \subseteq \mathbb{R}^n$ is an affine subspace if and only if, for every distinct $x, y \in \mathcal{A}$, the line $\{\lambda x + (1-\lambda)y\}$ passing through x and y belongs to $\mathcal{A}$. Furthermore a subset $\mathcal{A}$ of $\mathbb{R}^n$ is an affine space if and only if $\mathcal{A} = \{x \in \mathbb{R}^n \,|\, Ax = b\}$ for some matrix A. Note that the linear subspaces are precisely the affine subspaces containing the origin.

The *dimension* of a set $S \subseteq \mathbb{R}^n$, denoted by $\dim(S)$, is the maximum number of affinely independent points in S minus one. So the dimension of the empty set is -1, the dimension of a point is 0 and the dimension of a segment is 1. If $\mathcal{A} = \{x \in \mathbb{R}^n \,|\, Ax = b\}$ is nonempty, then $\dim(\mathcal{A}) = n - \text{rank}(A)$.

The inclusionwise minimal affine space containing a set $S \subseteq \mathbb{R}^n$ is called the *affine hull* of S and is denoted by $\text{aff}(S)$. Since the intersection of affine spaces is an affine space, the affine hull is well-defined. Note that $\dim(S) = \dim(\text{aff}(S))$.

Convex combinations, convex sets

A point x in $\mathbb{R}^n$ is a *convex combination* of the points $x^1, \ldots, x^q \in \mathbb{R}^n$ if there exist *nonnegative* scalars $\lambda_1, \ldots, \lambda_q$ such that

$$x = \sum_{j=1}^{q} \lambda_j x^j, \quad \sum_{j=1}^{q} \lambda_j = 1.$$

In particular, given three points x, y, z in $\mathbb{R}^n$, the point x is a *convex combination* of y and z if there exists $\lambda \in [0,1]$ such that $x = \lambda y + (1-\lambda)z$, that is, x is contained in the line segment joining y and z. If $y \neq z$ and $\lambda \in (0,1)$, then we say that x is a *proper convex combination* of y and z.

Definition 2.1. *A set $C \subseteq \mathbb{R}^n$ is* convex *if C contains all convex combinations of points in C. Equivalently, $C \subseteq \mathbb{R}^n$ is convex if for any two points $x, y \in C$, the line segment $\{\lambda x + (1-\lambda)y,\ 0 \leq \lambda \leq 1\}$ with endpoints x, y is contained in C.*

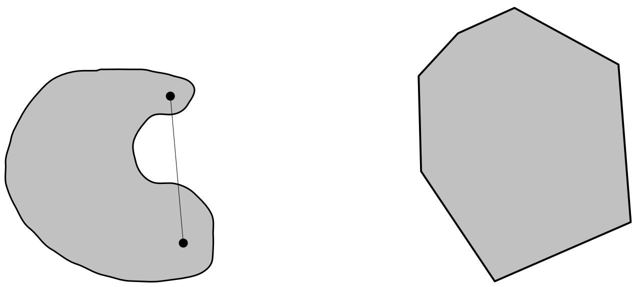


Figure 2.1: The set on the left is not convex, the one on the right is convex.

It is an easy exercise to show that every half-space is convex, and that the intersection of convex sets is convex. This shows that every polyhedron (and thus the feasible region of every LP problem) is a convex set.

Given a set $S \subseteq \mathbb{R}^n$, the *convex hull* of S, denoted by $\text{conv}(S)$, is the inclusionwise minimal convex set containing S. As the intersection of convex sets is a convex set, $\text{conv}(S)$ exists. Moreover, it is the set of all points that are convex combinations of points in S. That is

$$\text{conv}(S) = \left\{ \sum_{j=1}^{q} \lambda_j x^j \,|\, x^1, \ldots, x^q \in S,\ \lambda_1, \ldots, \lambda_q \geq 0,\ \sum_{j=1}^{q} \lambda_j = 1 \right\}.$$

Conic combinations, convex cones

A vector $x \in \mathbb{R}^n$ is a *conic combination* of vectors $x^1, \ldots, x^q \in \mathbb{R}^n$ if there exist scalars $\lambda_j \geq 0$, $j = 1, \ldots, q$, such that

$$x = \sum_{j=1}^{q} \lambda_j r^j.$$

A set $C \subseteq \mathbb{R}^n$ is a *cone* if $0 \in C$ and for every $x \in C$ and $\lambda \geq 0$, λx belongs to C. In other words, C is a cone if and only if $0 \in C$ and, for every $x \in C \setminus \{0\}$, C contains the half line starting from the origin in the direction x.

A cone C is a *convex cone* if C contains every conic combination of vectors in C. A convex cone is a convex set, since by definition every convex combination of points is also a conic combination.

Given a nonempty set $S \subseteq \mathbb{R}^n$, the *cone of* S, denoted by cone(S), is the inclusionwise minimal convex cone containing S. As the intersection of convex cones is a convex cone, cone(S) exists. It is the set of all conic combinations of vectors in S. We say that cone(S) is the cone generated by S. For convenience, we define cone$(\emptyset) := \{0\}$.

Given a cone C and a vector $r \in C \setminus \{0\}$, the half line cone$(r) = \{\lambda r,\ \lambda \geq 0\}$ is called a *ray* of C. We will often simply refer to a vector $r \in C \setminus \{0\}$ as a ray of C to denote the corresponding ray cone(r). Since cone$(\lambda r) =$ cone(r) for every $\lambda > 0$, we say that two rays r and r' of a cone are distinct when there is no $\mu > 0$ such that $r = \mu r'$.

The geometric interpretation of Farkas' lemma

Farkas' lemma has a nice geometric interpretation for cones. Let A be an $m \times n$ matrix. Consider the form as in Theorem 1.6: either $Ax = b$, $x \geq 0$ is feasible (*primal system*), or $u^\top A \leq 0$, $u^\top b > 0$ is feasible (*dual system*).

Let $A_1, A_2, \ldots, A_n \in \mathbb{R}^m$ denote the column vectors of the matrix A. We let $C = \text{cone}(\{A_1, A_2, \ldots, A_n\})$ denote the cone generated by these vectors, that is, the set of all conic combinations of these vectors. We can write

$$C = \{y \in \mathbb{R}^m \mid \exists x \in \mathbb{R}^n, x \geq 0, y = Ax\}.$$

Then, the primal system is feasible if and only if $b \in C$.

Let us now consider the dual system. The half-space $H = \{y \in \mathbb{R}^m \mid u^\top y \leq 0\}$ contains all vectors A_i, $i = 1, \ldots, m$, and consequently, the entire cone C. On the other hand, it does not contain b, thereby certifying infeasibility of $Ax = b$, $x \geq 0$. This is illustrated in Figure 2.2.

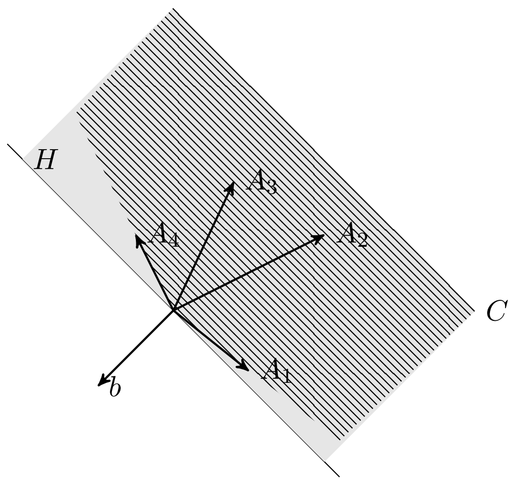


Figure 2.2: Illustration of Farkas' lemma

2.2 Valid inequalities, faces and facets

An inequality $c^\top x \le \delta$ is *valid* for the set $P \subseteq \mathbb{R}^n$ if $c^\top x \le \delta$ is satisfied by every point in P. Note that we allow $c = 0$ in our definition, in which case the inequality $0 \le \delta$ is valid for every set P if $\delta \ge 0$, and it is valid only for the empty set if $\delta < 0$.

Theorem 2.2. *Let $P := \{x \in \mathbb{R}^n \mid Ax \le b\}$ be a nonempty polyhedron. An inequality $c^\top x \le \delta$ is valid for P if and only if there exists $u \ge 0$ such that $u^\top A = c$ and $u^\top b \le \delta$.*

Proof. Let $c \in \mathbb{R}^n$ and $\delta \in \mathbb{R}$. Assume $u^\top A = c$, $u^\top b \le \delta$, $u \ge 0$ is feasible. Then, for all $x \in P$, we have $c^\top x = u^\top Ax \le u^\top b \le \delta$. This shows that $c^\top x \le \delta$ is valid for P.

Conversely, assume that the inequality $c^\top x \le \delta$ is valid for P. Consider the linear program $\max\{c^\top x \mid x \in P\}$. Since $P \ne \emptyset$ and $c^\top x \le \delta$ is a valid inequality for P, the above program admits a finite optimum and its value is $\delta' \le \delta$. By Proposition 1.9, the set $D = \{u \mid u^\top A = c,\ u \ge 0\}$ is nonempty. By Theorem 1.7 the dual $\min\{u^\top b \mid u \in D\}$ has value δ', and there exists $u \in D$ such that $u^\top b = \delta'$. Thus $u^\top A = c$, $u^\top b \le \delta$, $u \ge 0$. $\square$

A *face* of a polyhedron P is a set of the form

$$F := P \cap \{x \in \mathbb{R}^n \mid c^\top x = \delta\}$$

where $c^\top x \le \delta$ is a valid inequality for P. We say that the inequality $c^\top x \le \delta$ *defines* the face F. If a valid inequality $c^\top x \le \delta$ with $c \ne 0$ defines a nonempty face of P, the hyperplane $\{x \in \mathbb{R}^n \mid c^\top x = \delta\}$ is called a *supporting hyperplane* of P. A face is itself a polyhedron since it is the intersection of the polyhedron P with another polyhedron (the hyperplane $c^\top x = \delta$ when $c \ne 0$). Note that $\emptyset$ and P are always faces of P, since they are the faces defined by the valid inequalities $0 \le -1$ and $0 \le 0$, respectively. A face of P is said to be *proper* if it is nonempty and properly contained in P.

Inclusionwise maximal proper faces of P are called *facets*. Thus any face distinct from P is contained in a facet. Any valid inequality for P that defines a facet is called a *facet-defining inequality*.

Theorem 2.3 (Characterization of the faces)**.** *Let $P := \{x \in \mathbb{R}^n \mid a_i^\top x \le b_i,\ i \in M\}$ be a nonempty polyhedron. For any $I \subseteq M$, the set*

$$F_I := \{x \in \mathbb{R}^n \mid a_i^\top x = b_i, i \in I,\ a_i^\top x \le b_i, i \in M \setminus I\}$$

is a face of P. Conversely, if F is a nonempty face of P, then $F = F_I$ for some $I \subseteq M$.

Proof. For the first part of the statement, let $c := \sum_{i\in I} a_i$, $\delta := \sum_{i\in I} b_i$. Then $c^\top x \le \delta$ is a valid inequality. Furthermore, given $x \in P$, x satisfies $c^\top x = \delta$ if and only if it satisfies $a_i^\top x = b_i$, $i \in I$. Thus $F_I = P \cap \{x \in \mathbb{R}^n \mid c^\top x = \delta\}$, so F_I is a face.

Conversely, let $F := \{x \in P : c^\top x = \delta\}$ be a nonempty face of P defined by the valid inequality $c^\top x \le \delta$. Then F is the set of optimal solutions of the linear program $\max\{c^\top x \mid x \in P\}$. Let $\bar{u}$ be an optimal solution to the dual problem $\min\{u^\top b \mid u^\top A = c,\ u \ge 0\}$, and $I = \{i \in M \mid \bar{u}_i > 0\}$. By Theorem 1.8, $F = F_I$. $\square$

2.3 Vertices and basic solutions

The 2-dimensional examples seen in MA423 suggest that the optimum of an LP is always attained by some vertex of the feasible region. This notion can be generalized to problems in arbitrary dimension (i.e. with arbitrarily many variables). We first need to give a good definition of what we mean by a "vertex". We consider three different geometric definitions, and then show that the three notions coincide.

First definition The first intuition is that a vertex is a face of dimension 0. That is, for a polyhedron P and valid inequality $c^\top x \leq \delta$, the intersection $P \cap \{c^\top x = \delta\}$ is a single point x^*. This means that there is some hyperplane tangent to P that only touches that point (see Figure 2.3 to the left). We formalize this concept in the following definition.

Definition 2.4 (Vertex)**.** *Given a polyhedron $P \subseteq \mathbb{R}^n$, $x^* \in P$ is a* vertex *of P if it is a face of dimension 0. That is, there exists some $c \in \mathbb{R}^n$ such that $c^\top x < c^\top x^*$ for every $x \in P \setminus \{x^*\}$.*

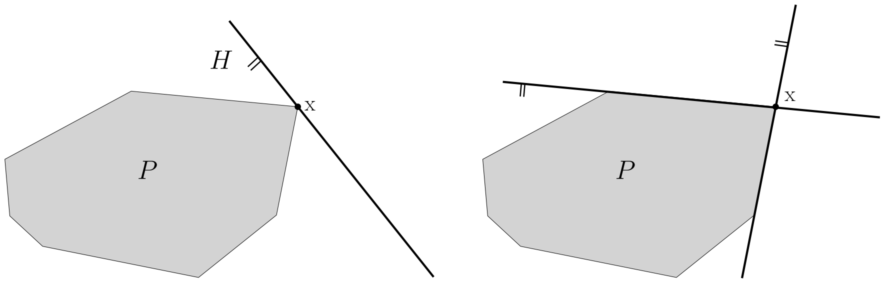


Figure 2.3: Two explanations of why point x^* is a vertex of the polyhedron P.

Second definition The intuition that is perhaps most natural is that a vertex of a polyhedron P in n dimension is a point that lies on the intersection of the boundaries of n of the half-spaces that define P (see Figure 2.3 to the right).

This notion is a bit misleading, however. For example, in Figure 2.4, the point x^* is not a vertex of the 3-dimensional polyhedron depicted (to clarify the picture, the polyhedron is the one lying under the two gray planes, whereas the blue plane does represents a redundant constraint), however it is at the intersection of the boundaries of three half-spaces. The problem is that one of the three half-spaces (the one defined by the plane colored in dark blue) is not needed to define the polyhedron, in the sense that it does not cut off anything that is not already cut off by the other two half-spaces.

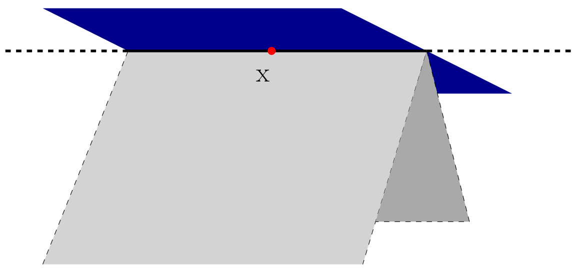


Figure 2.4: The point x^* is not a vertex of the 3-dimensional polyhedron even though it lies on the intersection of three half-planes defining P.

We need therefore to be more careful in our definition. Given constraints $a_i^\top x = b_i$, $i = 1, \ldots, k$, $a_i^\top x \leq b_i$, $i = k+1, \ldots, m$, we say that the constraints are *linearly independent* if the vectors $\{a_i, | \, i = 1, \ldots, m\}$ are linearly independent. Given a point $x^* \in \mathbb{R}^n$, we say that the ith constraint $(i = 1, \ldots, n)$ is *binding* at x^* if $a_i^\top x^* = b_i$ (note that equality constraints are always binding for every feasible point).

Definition 2.5 (Basic feasible solution)**.** *Given a system of linear constraints in $\mathbb{R}^n$, a point $x^* \in \mathbb{R}^n$ is said a* basic feasible solution *of the system if it is feasible and there exist n linearly independent constraints of the system that are binding at x^*.*

Third definition Our third definition is in terms of convex combinations. Recall that every polyhedron P is a convex set. We define extreme points more generally, for arbitrary convex sets. These are the "corner" points, that are not included inside line segments between other points.

Definition 2.6 (Extreme points). *Given a convex set S, we say that $x^* \in P$ is an* extreme point in S, *if x^* is not a proper convex combination of two distinct points in S.*

The next theorem reconciles these three apparently different intuitions by stating that they are indeed equivalent, so any of them will capture the correct notion of a vertex.

Theorem 2.7. *Let $P = \{x \mid Ax \leq b\}$ be a polyhedron in $\mathbb{R}^n$, where $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$. Given a point $x^* \in P$, the following are equivalent.*

(i) x^* is a vertex of P.

(ii) x^* is a basic feasible solution of the system $Ax \leq b$.

(iii) x^* is an extreme point of P.

Proof. We first show that **(i)** implies **(iii)**: *every vertex must be an extreme point.* Indeed, suppose x^* is a vertex, and let $c^\top x \leq \delta$ be a valid inequality for P such that $x^* = P \cap \{x \mid c^\top x = \delta\}$. Given $x', x'' \in P$ and $0 < \lambda < 1$, such that $x^* = \lambda x' + (1-\lambda)x''$, it follows that $\delta = c^\top x^* = \lambda c x' + (1-\lambda)c x'' \leq \delta$, therefore $c^\top x' = c^\top x'' = \delta$. Byt the assumption $x^* = P \cap \{x \mid c^\top x = \delta\}$, we get that $x' = x'' = x^*$. Hence, x^* cannot be written as a proper convex combination of two distinct points in P.

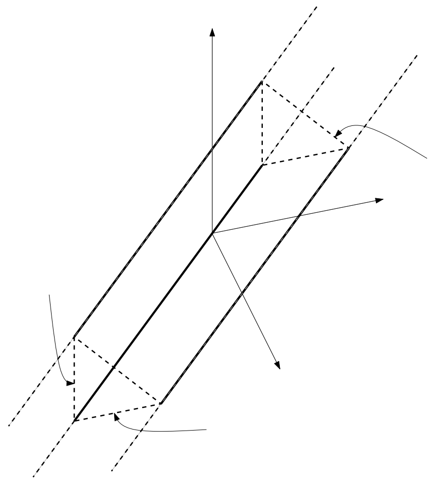
Next, we show that **(iii)** implies **(ii)**: *every extreme point must be basic feasible* . Assume that x^* is an extreme point, and let $A'x \leq \bar{b}'$ be the system comprising all inequalities of $Ax \leq b$ satisfied at equality by x^*. We will show that, if $\text{rank}(A') < n$, then x^* is a proper convex combination of two points in P, which gives a contradiction. If $\text{rank}(A') < n$, then there exists a vector $y \neq 0$ such that $A'y = 0$. Let $A''x \leq b''$ be the system of inequalities in $Ax \leq b$ satisfied with strict inequality for x^*. Thus, for $\varepsilon > 0$ sufficiently small, the points $x' = x^* + \varepsilon y$ and $x'' = x^* - \varepsilon y$ are both in P. It is now clear that $x^* \neq x', x''$ and $x^* = \frac{1}{2}x' + \frac{1}{2}x''$. This contradicts the assumptions that x^* was an extreme point.

It remains to show that **(ii)** implies **(i)**: *every basic feasible solution is a vertex.* Let x^* be a basic feasible solution, and again let $A'x \leq b'$ be the system comprising all inequalities of $Ax \leq b$ satisfied at equality by x^*. Thus, $\text{rank}(A') = n$. Let us define $c := \sum_i a'_i$, and $\delta := \sum_i b'_i$. That is, c is the sum of all right hand side vectors of all binding inequalities at x^*, and δ is the sum of the left hand side values. We claim that $x^* = P \cap \{x \mid c^\top x = \delta\}$, showing that x^* is a vertex. If $c^\top x = \delta$ for some $x \in P$, then $A'x = b'$ must hold, since $c^\top x \leq \delta$ is the sum of all inequalities in the system $A'x \leq b'$. Therefore, if any inequality here was strict, then $c^\top x < \delta$ would follow. However, x^* is the unique solution of the system of linear equations $A'x = b'$, since $\text{rank}(A') = n$. $\square$

NOTE: In general, a system of linear constraints might not have any basic feasible solution. According to the Theorem above, such polyhedra cannot have any vertices nor extreme points. For example, consider the system

$$\begin{aligned} x_1 + x_2 + x_3 &\leq 1, \\ x_1 + x_2 \phantom{{}+x_3} &\geq 0, \\ x_3 &\geq 0, \end{aligned}$$

The polyhedron defined by these inequalities is depicted below (dashed lines are meant to represent the fact that the polyhedron extends indefinitely).



Note that the system is in three variables and is comprised only of 3 constraints. However, the three constraint are not linearly independent, since $(1,1,1)-(1,1,0)-(0,0,1)=0$. Hence the system cannot have a basic feasible solution.

However, whenever a feasible system of linear constraints contains n linearly independent constraints, then it has a basic solution, as the next theorem states.

Theorem 2.8. *Let $P = \{x \mid Ax \leq b\}$ be a polyhedron in $\mathbb{R}^n$, where $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$. Assume that the problem is feasible, that is, $P \neq \emptyset$. Then, the following three properties are equivalent:*

(i) There exists a basic feasible solution in P.

(ii) The matrix A has $\mathrm{rk}(A) = n$.

(iii) The only solution of $Az = 0$ is $z = 0$.

We do not prove this theorem here. However, let us note that the equivalence of (ii) and (iii) is immediate: A has rank n, and therefore the columns are linearly independent, which is the same as (iii).

In particular, any feasible system of linear constraints in which all variables are restricted to be nonnegative has a basic feasible solution, since the n constraints $x_j \geq 0$, $j = 1,\ldots,n$ are linearly independent, showing that the rank of the constraint matrix is n. This is the case for LPs in standard form, that is, $Ax \leq b$, $x \geq 0$. This can be written in the standard inequality form as $\binom{A}{-I}x \leq \binom{b}{0}$. Similarly, we obtain the existence of a basic feasible solution for the standard equality form, $Ax = b$, $x \geq 0$.

Furthermore, the following important result holds. We will prove the equivalent Theorem 2.16 in the next section.

Theorem 2.9. *If a linear programming problem that has a basic feasible solution admits an optimal solution, then there exists at least one optimal solution which is a basic feasible solution.*

By the above theorem, in order to find an optimal solution one only needs to consider the basic feasible solutions. Since a system of linear constraints has a finite number of basic solutions (indeed, a system of m constraints in n variables can have at most as many basic feasible solutions as the number of subsets of $\{1,\ldots,m\}$ with n elements, that is, $\binom{m}{n}$), the above theorem allows us to find an optimum by exploring only a finite number of possible points. However, the number of basic solutions, although finite, could be extremely large, so in general one cannot try all possibilities. The Simplex Method can be seen as a “clever” way of searching an optimal solution among the basic feasible solutions.

2.4 Standard equality form

The Simplex Method is applied to LP problems in standard equality form. Recall that any LP problem can be transformed into an equivalent problem in standard equality form, therefore in the remainder of the notes we will focus on such problems. Let P be a polyhedron defined by

$$P := \{x \in \mathbb{R}^n \,|\, Ax = b,\, x \geq 0\},$$

where $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$. We assume that P is nonempty, that is, we assume that the system $Ax \leq b$, $x \geq 0$ has a feasible solution. In particular, the system of equations $Ax = b$ has a solution.

Recall that $Ax = b$ has a solution if and only if $\mathrm{rk}(A) = \mathrm{rk}(A|b)$. We may assume therefore assume that $\mathrm{rk}(A) = \mathrm{rk}(A|b)$, otherwise the system is infeasible.

Recall also that, if $\mathrm{rk}(A) = \mathrm{rk}(A|b)$ then if we take a subsystem of $Ax = b$, say $A'x = b'$, where the matrix A' is comprised of $\mathrm{rk}(A)$ linearly independent rows of A, then the systems $Ax = b$ and $A'x = b'$ are equivalent, in the sense that they have the same solutions. In particular, we could replace $Ax = b$ with $A'x = b'$ in the definition of P without changing the polyhedron, so we may assume that A has *full row rank*, that is $\mathrm{rk}(A)$ equals the number of rows of A. Therefore, throughout the notes, we will always make the following assumption.

Assumption: $m = \mathrm{rk}(A)$, that is, the rows of A are linearly independent.

Recall that A_j denotes the j'th column of A. Given any set $S \subseteq \{1,\dots,n\}$, we will denote by A_S the submatrix of A formed by the columns of A with index in S, that is A_j, $j \in S$.

Definition 2.10. *A set $B \subseteq \{1,\dots,n\}$ is said a* basis *of A if*

- $|B| = m$;
- *the vectors A_j, $j \in B$, are linearly independent.*

Notice that, under the assumption that $m = \mathrm{rk}(A)$, there must exist at least one basis of A. Note that a subset B of $\{1,\dots,n\}$ is a basis if the matrix A_B is square and nonsingular. In particular, note that if B is a basis of A, then the system $A_B z = b$ has a unique solution (where z is a vector of m variables).

Proposition 2.11. *A point $x^* \in \mathbb{R}^n$ is a basic feasible solution of $Ax = b$, $x \geq 0$ if and only if it is feasible and there exists a basis B of A such that $x^*_j = 0$ for every $j \notin B$.*

Proof: The feasible point x^* is a basic solution of $Ax = b$, $x \geq 0$, if and only if there are n linearly independent inequalities binding at x^*. Since all m constraints $Ax = b$ are binding at x^*, and they are linearly independent, we need to identify a subset N of $n-m$ nonnegativity constraints $x_j \geq 0$, $j \in N$, that are binding at x^* and such that the n constraints $Ax = b$, $x_j \geq 0$, $j \in N$ are linearly independent. Assuming, up to reordering the variables, that $N = \{m+1,\dots,n\}$, and letting $B = \{1,\dots,m\}$, we have that the constraint matrix R of the system $Ax = b$, $x_j \geq 0$, $j \in N$, is of the following form

$$R = \left(\begin{array}{c|c} A_B & A_N \\ \hline \mathbf{0} & I \end{array}\right), \tag{2.1}$$

where I is the $(n-m)\times(n-m)$ identity matrix. From standard linear algebra, the rows of R are linearly independent if and only if $\det(R) \neq 0$. Also, $\det(R) = \det(A_B)$. Thus x^* is a basic feasible solution if and only if $\det(A_B) \neq 0$, that is, if and only if B is a basis. Therefore x^* is a basic solution of $Ax = b$, $x \geq 0$, if and only if $Ax^* = b$ and there exists some basis B such that $x^*_j = 0$ for all $j \notin B$. □

Later on in the notes, when studying the dual Simplex Method, we will need to have a notion of basic solutions that are not necessarily feasible. This motivates the following definition, which is consistent with Proposition 2.11.

Definition 2.12. *A point $x^* \in \mathbb{R}^n$ is a basic solution of $Ax = b$, $x \geq 0$ if $Ax^* = b$ and there exists a basis B of A such that $x^*_j = 0$ for every $j \notin B$.*

Corollary 2.13. *A point $x^* \in \mathbb{R}^n$ is a basic solution of $Ax = b$, $x \geq 0$ if and only if $Ax^* = b$ and the vectors in $\{A_j \,|\, x^*_j \neq 0\}$ are linearly independent.*

Proof: If x^* is a basic solution, then by the above proposition there exists a basis B of $Ax = b$, $x \geq 0$ such that $x^*_j = 0$ for all $j \notin B$. Hence $\{A_j \,|\, x^*_j \neq 0\} \subseteq \{A_j \,|\, j \in B\}$, and all these vectors are linearly independent by definition of a basis.

Viceversa, assume the vectors in $\{A_j \,|\, x^*_j \neq 0\}$ are linearly independent. Then, by a standard linear algebra fact, we can choose a bases B of $Ax = b$, $x \geq 0$ such that $\{j \,|\, x^*_j \neq 0\} \subseteq B$. Therefore $x^*_j = 0$ for every $j \notin B$, so by the previous proposition x^* is basic. □

Lemma 2.14. *Given a basis B of $Ax = b$, $x \geq 0$, there exists a unique vector x^* such that $Ax^* = b$, $x^*_j = 0$, $j \notin B$.*

Proof: Up to permuting variables and columns of A accordingly, we may assume that $B = \{1, \ldots, m\}$. Let $N = \{m+1, \ldots, n\}$. Let us split the vector x into two vectors: the vector x_B formed by the first m components of x (i.e. $x_B = (x_1, \ldots, x_m)^\top$), and the vector x_N formed by the remaining $n - m$ components (i.e. $x_N = (x_{m+1}, \ldots, x_n)^\top$). Hence $x = (x_B, x_N)^\top$.

Any vector x that satisfies $Ax = b$ and $x_j = 0$ for $j \in N$ must satisfy $A_B x_B = b$, $x_N = 0$. Since B is a basis, A_B is a square nonsingular matrix, therefore there exists a unique vector satisfying $A_B x_B = b$, namely the vector $A_B^{-1} b$. Hence there exists a unique vector x^* such that $Ax^* = b$, $x^*_j = 0$, $j \notin B$, namely $x^* = (A_B^{-1} b, 0)^\top$. □

Lemma 2.14 says that *every basis* has a *unique basic solution* associated to it. Formally:

Definition 2.15. *Given a basis B of $Ax = b$, $x \geq 0$, the* basic solution relative to B *is the unique vector x^* such that $Ax^* = b$, $x^*_j = 0$, $j \notin B$; namely, the vector*

$$\begin{aligned} x^*_B &= A_B^{-1} b; \\ x^*_N &= 0 \end{aligned} \tag{2.2}$$

If x^* is feasible (i.e., if $A_B^{-1} b \geq 0$), B is said a feasible basis *of $Ax = b$, $x \geq 0$.*

Example Consider the following matrix A and right-hand-side b.

$$A = \begin{bmatrix} 1 & 2 & 0 & 1 & 0 & -6 \\ 0 & 1 & 1 & 3 & -2 & -1 \\ 1 & 2 & 1 & 3 & -1 & -5 \end{bmatrix}, \quad b = \begin{bmatrix} 11 \\ 6 \\ 13 \end{bmatrix}.$$

Consider the set $B = \{2, 5, 6\}$. It can be verified that B is not a basis of $Ax = b$, $x \geq 0$, since $\det(A_B) = 0$. Another way to verify this is to observe that $(3, 1, 1)^\top$ is a nonzero solution to the system $A_B z = 0$.

Consider the set $B = \{2, 3, 6\}$. It can be verified that B is a basis of $Ax = b$, $x \geq 0$, since $\det(A_B) = -2$. We can compute

$$A_B^{-1} = \begin{bmatrix} 2 & 3 & -3 \\ -\frac{3}{2} & -1 & 2 \\ \frac{1}{2} & 1 & -1 \end{bmatrix}.$$

We have then

$$\bar{x}_B = \begin{bmatrix} \bar{x}_2 \\ \bar{x}_3 \\ \bar{x}_6 \end{bmatrix} = A_B^{-1} b = \begin{bmatrix} 1 \\ \frac{7}{2} \\ -\frac{3}{2} \end{bmatrix}.$$

The basic solution associated to the basis B is therefore $\bar{x} = (0, 1, \frac{7}{2}, 0, 0, -\frac{3}{2})^\top$. Note that $\bar{x}$ is not feasible, because $\bar{x}_6 < 0$. Hence B is not a feasible basis.

Consider now the set $B = \{1, 2, 3\}$. It can be verified that B is a basis of $Ax = b$, $x \geq 0$, since $\det(A_B) = 1$. We can compute

$$A_B^{-1} = \begin{bmatrix} -1 & -2 & 2 \\ 1 & 1 & -1 \\ -1 & 0 & 1 \end{bmatrix}.$$

We have then

$$\bar{x}_B = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \bar{x}_3 \end{bmatrix} = A_B^{-1} b = \begin{bmatrix} 3 \\ 4 \\ 2 \end{bmatrix}.$$

The basic solution associated to the basis B is therefore $\bar{x} = (3, 4, 2, 0, 0, 0)^\top$. Note that $\bar{x}$ is feasible because $\bar{x} \geq 0$. Hence B is a feasible basis.

2.5 Existence of basic feasible solutions

The following theorem is the main result of this chapter, and it is a fundamental result in Linear Programming. Note that the theorem is precisely Theorem 2.9 restated for LP problems in standard form.

Theorem 2.16. *If the linear programming problem*

$$\begin{aligned} \max c^\top x & \\ Ax = & \; b \\ x \geq & \; 0 \end{aligned} \tag{2.3}$$

has a finite optimum, there exists an optimal solution x^* which is a basic feasible solution.

The Simplex Method will provide a proof, since it terminates with a basic optimal solution whenever an optimal solution exists. We now give a direct proof.

Proof: Assume that (2.3) has a finite optimum. The problem might have more than one optimum, and some of them might not be basic. We need to show that there exists one optimal solution that is a basic feasible one. Among all optimal solutions of (2.3), choose one that has the highest number of components at zero. Let us a call such an optimal solution x^*. We will show that x^* is basic, thus proving the theorem.

Suppose, by contradiction, that x^* is not a basic solution. To derive a contradiction, we will show that there must exist some other optimal solution that has more components at zero than x^*, contrary to our choice of x^*.

Let S be the set of variable indices corresponding to the positive components of x^*, i.e. $S = \{j \mid x_j^* > 0\}$. By Corollary 2.13, the vectors A_j, $j \in S$ cannot be linearly independent, otherwise x^* would be basic. Thus, we can write the null vector as a nonzero linear combination of the vectors A_j, $j \in S$. In other words, there exists multipliers z_j, $j \in S$, not all zero, such that $\sum_{j \in S} A_j z_j = 0$.

Define a vector $d \in \mathbb{R}^n$ as follows. For $j = 1, \ldots, n$, let

$$d_j = \begin{cases} z_j & \text{if } j \in S \\ 0 & \text{if } j \notin S \end{cases}$$

Observe that, by the way we defined d, $Ad = 0$, $d \neq 0$, and $d_j = 0$ for every $j \notin S$.

Observe now that, for every scalar $t \geq 0$, we have that $A(x^* - td) = Ax^* - tAd = b$ and $x^* - td_j = 0$ for every $j \notin S$. Thus, for $\epsilon > 0$ sufficiently small, $x^* + \epsilon d$ and $x^* - \epsilon d$ are both feasible solutions. Furthermore, $c^\top(x^* + \epsilon d) = c^\top x^* + \epsilon c^\top d$ and $c^\top(x^* - \epsilon d) = c^\top x^* - \epsilon c^\top d$. Thus $c^\top d = 0$, otherwise either $x^* + \epsilon d$ or $x^* - \epsilon d$ would be a feasible solution with higher objective value than x^*, contradicting the fact that x^* is optimal. In particular, $x^* - td$ is an optimal solution for every $t \in \mathbb{R}$ such that $(x^* - td)$ is feasible.

We may assume without loss of generality that d has at least one positive component, otherwise we can just redefine d as $-d$.

Let $\bar{t} > 0$ be the largest possible scalar such that $\bar{x} = x^* - \bar{t}d$ is feasible for (2.3). This point is feasible if and only if $x^* - \bar{t}d \geq 0$. This happens if and only if

$$\bar{t} \leq \frac{x_j^*}{d_j}, \quad \text{for every } j \in \{1, \ldots, n\} \text{ such that } d_j > 0.$$

Hence $\bar{t} = \min\{\frac{x^*_j}{d_j} \mid j \in \{1, \ldots, n\}$ such that $d_j > 0\}$. Since d has at least one positive component, $\bar{t}$ is some finite number. Note also that, for every index h such that $d_h > 0$, by construction we have that also $x^*_h > 0$. Hence $\bar{t} > 0$, and there exists some index h such that $d_h > 0$ and $\bar{t} = \frac{x^*_h}{d_h}$.

But then, $\bar{x}_h = x^*_h - \bar{t}d_h = x^*_h - \frac{x^*_h}{d_h}d_h = 0$, and $\bar{x}_j = x^*_j + \bar{t}d_j = 0+0 = 0$ for every $j \notin S$. This show that $\bar{x}$ is an optimal solution that has at least one more component at zero than x^*, a contradiction. □

The next corollary states that, unlike in the general case where a polyhedron might not have any vertices, polyhedra defined by systems of constraints in standard equality form always have vertices.

Corollary 2.17. *If $Ax = b$, $x \geq 0$ has a feasible solution, then it has a basic feasible solution.*

Proof: Consider the linear programming problem

$$\begin{aligned} \max 0^\top x & \\ Ax = \ & b \ . \\ x \geq \ & 0 \end{aligned}$$

Clearly every feasible solution to the above is also optimal, since the objective function is constantly 0. By Theorem 2.16, the problem (2.3) has an optimal solution that is a basic feasible solution. In particular, the system $Ax = b$, $x \geq 0$, has a basic feasible solution. □

2.6 Carathéodory's theorem

We now exhibit classical results from geometry which can be easily derived using the concept of basic feasible solutions, in particular, from Corollary 2.17.

Theorem 2.18 (Carathéodory). *If a point $z \in \mathbb{R}^n$ is a convex combination of points in some set $X \subseteq \mathbb{R}^n$, then it is a convex combination of at most* $\dim(X) + 1$ *affinely independent points in X.*

Proof. We can assume that X is finite, say $X = \{v^1, v^2, \ldots, v^k\}$. If $z \in \operatorname{conv}(X)$, then there exists a vector $\lambda \in \mathbb{R}^k$, $\lambda \geq 0$ such that $z = \sum_{i=1}^k \lambda_i v^i$, and $\sum_{i=1}^k \lambda_i = 1$.

Let us construct the matrix $A \in \mathbb{R}^{(n+1)\times k}$ as follows: let the ith column be the vector $\binom{v^i}{1} \in \mathbb{R}^{n+1}$, that is, the n dimensional vector v^i, with an entry 1 appended. Then, λ represents the coefficients in a convex combination giving z if and only if it is a solution to the system

$$\begin{aligned} A\lambda &= z \\ \lambda &\geq 0. \end{aligned}$$

By Corollary 2.17, this system has a basic feasible solution. From now on, let us assume that λ is selected as a basic feasible solution. Recall from Proposition 2.11 that this corresponds to a basis $B \subseteq \{1, 2, \ldots, k\}$ such that $\lambda_j = 0$ for every $j \notin B$, and the columns A^j of A corresponding to indices $j \in B$ are linearly independent. For simplicity of notation, let us assume $B = \{1, 2, \ldots, n+1\}$ (we can always renumber the columns to achieve this). In other words, the vectors $\binom{v^1}{1}, \ldots, \binom{v^{n+1}}{1}$ are linearly independent, and $\sum_{i=1}^{n+1} \lambda_i v^i = z$, $\sum_{i=1}^{n+1} \lambda_i = 1$, $\lambda \geq 0$.

The linear independence means that there is no vector $y \in \mathbb{R}^{n+1}$ such that $\sum_{i=1}^{n+1} y_i v^i = 0$ and $\sum_{i=1}^{n+1} y_i = 0$, which is the same as saying that the vectors $v^1, \ldots, v^{n+1}$ are affinely independent. The statement follows. □

We now state two more theorems; their proofs are left as exercises.

Theorem 2.19 (Radon). *Let S be a subset of $\mathbb{R}^d$ with at least $d+2$ points. Then S can be partitioned into two sets S_1 and S_2 so that* $\operatorname{conv}(S_1) \cap \operatorname{conv}(S_2) \neq \emptyset$.

Theorem 2.20 (Helly). *Let $C_1, C_2, \ldots, C_h$ be convex sets in $\mathbb{R}^d$ such that $C_1 \cap C_2 \cap \cdots \cap C_h = \emptyset$, where $h \geq d+1$. Then there exist $d+1$ sets among $C_1, C_2, \ldots, C_h$ whose intersection is empty.*

Chapter 3

The Simplex Method

3.1 The tableau form

The Simplex Method is always applied to problems in standard form. (Note that LP solvers accept LP problems in any form, and they turn them into standard form by applying the transformations we described in the MA423 lecture notes.)

Consider then an LP problem in standard form

$$\begin{aligned} \max c^\top x & \\ Ax = & \ b \\ x \geq & \ 0 \end{aligned} \tag{3.1}$$

where $A \in \mathbb{R}^{m\times n}$, $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, and x is a vector of indeterminates in $\mathbb{R}^n$. We will adhere to the assumption made in the previous chapter: $m = \mathrm{rk}(A)$, that is, the rows of A are linearly independent.

Throughout this chapter, it will be convenient to write (3.1) in the following equivalent form:

$$\begin{aligned} \max z & \\ z - c^\top x = & \ 0 \\ Ax = & \ b \\ x \geq & \ 0 \end{aligned} \tag{3.2}$$

Let $B \subseteq \{1,\dots,n\}$ be a **feasible basis** of $Ax = b$, $x \geq 0$. A vector $x \in \mathbb{R}^n$ satisfies the above inequalities if and only if it satisfies

$$\begin{array}{rcl} A_B^{-1}Ax & = & A_B^{-1}b \\ x & \geq & 0 \end{array}.$$

For simplicity, let us assume for now that $B = \{1,\dots,m\}$, and thus $A = (A_B, A_N)$. We split x and c accordingly as $x = \binom{x_B}{x_N}$ and $c = \binom{c_B}{c_N}$. It follows that a vector $x \geq 0$ is feasible for (3.1) if and only if it satisfies

$$x_B = A_B^{-1}b - A_B^{-1}A_N x_N.$$

Substituting $A_B^{-1}b - A_B^{-1}A_N x_N$ for x_B in the equation $z - c_B^\top x_B - c_N^\top x_N = 0$, and bringing constant terms to the right-hand-side, we get

$$z - (c_N^\top - c_B^\top A_B^{-1} A_N)x_N = c_B^\top A_B^{-1} b.$$

Let us define

$$\begin{array}{rcl} \bar{A}_N & = & A_B^{-1}A_N; \\ \bar{b} & = & A_B^{-1}b; \\ \bar{c} & = & c - A^\top {A_B^{-1}}^\top c_B; \\ \bar{z} & = & c_B^\top A_B^{-1} b. \end{array} \tag{3.3}$$

The vector $\bar{c}$ is said the *vector of the reduced costs* of (3.1) relative to the basis B. Notice that

$$\begin{aligned} \bar{c}_B &= 0 \\ \bar{c}_N &= c_N - A_N^\top {A_B^{-1}}^\top c_B. \end{aligned} \tag{3.4}$$

It follows from the above discussion that (3.2) is equivalent to the following LP problem.

$$\begin{array}{lllcl} \max & z & & & \\ & z & -\bar{c}_N x_N & = & \bar{z} \\ & x_B & +\bar{A}_N x_N & = & \bar{b} \\ & & x & \geq & 0 \end{array} \tag{3.5}$$

Problem (3.5) is said to be *in tableau form with respect to the basis B*. We will often represent the above problem in a compact form through the matrix

$$\begin{array}{|c|cc|c|} \hline 1 & \mathbf{0} & -\bar{c}_N^\top & \bar{z} \\ \hline & & & \\ \mathbf{0} & I & \bar{A}_N & \bar{b} \\ \hline \end{array}\ ,$$

which is called the *tableau* of the problem (3.1) with respect to the basis B.

Remark 3.1. *In general, the basic variables will not be the first m, and they will not appear in order. Suppose, then, that the elements of B are $B[1], \ldots, B[m] \in \{1, \ldots, n\}$. So $A_B = [A_{B[1]}, \ldots, A_{B[m]}]$ and the problem in tableau form will be*

$$\begin{array}{lllcl} \max & z & & & \\ & z & -\sum_{j\in N} \bar{c}_j x_j & = & \bar{z} \\ & x_{B[i]} & +\sum_{j\in N} \bar{a}_{ij} x_j & = & \bar{b}_i, \quad i = 1, \ldots, m \\ & & x & \geq & 0. \end{array}$$

We will say that $x_{B[i]}$ is the basic variable in row i, $i = 1, \ldots, m$.

Note that the basic solution $\bar{x}$ with respect to B is defined by

$$\begin{aligned} \bar{x}_{B[i]} &= \bar{b}_i, \quad i = 1, \ldots, m \\ \bar{x}_j &= 0, \quad j \in N \end{aligned}$$

so in particular $\bar{b} \geq 0$ because B is a feasible basis. Furthermore, observe that the objective value of $\bar{x}$

$$\bar{z} + \sum_{j\in N} \bar{c}_j \bar{x}_j = \bar{z}.$$

Suppose that $\bar{c}_N \leq 0$. Note that, for every feasible solution $\tilde{x}$ of (3.1), the objective value of $\tilde{x}$ is

$$\bar{z} + \sum_{j\in N} \underbrace{\bar{c}_j}_{\leq 0} \underbrace{\tilde{x}_j}_{\geq 0} \leq \bar{z}.$$

In other words, if $\bar{c}_N \leq 0$, then $\bar{x}$ is an optimal solution for (3.1). We restate this fact in the next lemma.

Lemma 3.2. *If B is a feasible basis for (3.1) and the reduced costs of (3.1) relative to B are nonpositive, then the basic feasible solution relative to B is optimal.*

The above provides an optimality criterion: if we have determined a feasible basis whose reduced costs are nonpositive, then we have found a **provably optimal solution**. The aim of the simplex algorithm is to find such a basis.

Suppose then that $\bar{c}_k > 0$ for some $k \in N$. Ideally, we would like to find a new feasible basis whose objective value is better (or not worse) then $\bar{z}$. Note that, since $\bar{c}_k > 0$, if we increase the value of x_k

from 0 to some number $t \geq 0$, while leaving at 0 the value of all other nonbasic variables, then the value of the new solution in the objective function will be $\bar{z} + \bar{c}_j t \geq \bar{z}$. We would like to increase the value of x_k as much as possible. More formally, for $t \geq 0$, let $x(t)$ be the solution defined by

$$\begin{aligned} x_k(t) &= t; \\ x_{B[i]}(t) &= \bar{x}_{B[i]} - t\bar{a}_{ik} = \bar{b}_i - t\bar{a}_{ik}, \quad i = 1, \ldots, m; \\ x_j(t) &= 0, \quad j \in N \setminus \{k\}. \end{aligned} \tag{3.6}$$

By construction, $x(t)$ satisfies $x_{B[i]} + \sum_{j \in N} \bar{a}_{ij} x_j = \bar{b}_i$, $i = 1, \ldots, m$, therefore $Ax(t) = b$. Furthermore, the objective value of $x(t)$ is

$$\bar{z} + \sum_{j \in N} \bar{c}_j x_j(t) = \bar{z} + \bar{c}_j t \geq \bar{z}.$$

That is, the objective value of $x(t)$ is greater than or equal to the objective value of $\bar{x}$, where the value is strictly greater if and only if $t > 0$.

What is the largest possible value of t such that $x(t)$ is feasible? Clearly, since $x(t)$ satisfies the equality constraints for every choice of t, the vector $x(t)$ is feasible if and only if $x(t) \geq 0$. From 3.6, we get that $x(t)$ is feasible if and only if

$$\bar{b}_i - t\bar{a}_{ik} \geq 0, \quad i = 1, \ldots, m.$$

Let $i \in \{1, \ldots, m\}$. We have to possible cases.

1. $\bar{a}_{ik} > 0$: in this case $\bar{b}_i - t\bar{a}_{ik} \geq 0$ if and only if $t \leq \frac{\bar{b}_i}{\bar{a}_{ik}}$.

2. $\bar{a}_{ik} \leq 0$: in this case $\bar{b}_i - t\bar{a}_{ik} \geq 0$ for every choice of $t \geq 0$.

Hence $x(t)$ is feasible if and only if

$$t \leq \frac{\bar{b}_i}{\bar{a}_{ik}} \text{ for every } i \in \{1, \ldots, m\} \text{ such that } \bar{a}_{ik} > 0. \tag{3.7}$$

We have to consider two possible cases:

Case 1 There exists $i \in \{1, \ldots, m\}$ such that $\bar{a}_{ik} > 0$.

In this case, it follows from (3.7) that the largest $\bar{t}$ such that $x(\bar{t})$ is feasible is

$$\bar{t} = \min_{i \in \{1,\ldots,m\} : \bar{a}_{ik} > 0} \left\{ \frac{\bar{b}_i}{\bar{a}_{ik}} \right\}. \tag{3.8}$$

The above is called the *minimum quotient criterion* .

The new solution we compute is therefore $\tilde{x} = x(\bar{t})$. Let $h \in \{1, \ldots, m\}$ such that $\bar{a}_{hk} > 0$ and $\frac{\bar{b}_h}{a_{hk}} = \bar{t}$. We get that,

$$\tilde{x}_{B[h]} = \bar{b}_h - \bar{t}\bar{a}_{hk} = \bar{b}_h - \frac{\bar{b}_h}{\bar{a}_{hk}} \bar{a}_{hk} = 0,$$

hence, by construction,

$$\tilde{x}_j = 0, \quad \forall j \notin (B \cup \{k\}) \setminus \{B[h]\}.$$

Thus $\tilde{x}$ is the basic feasible solution relative to the basis $\tilde{B} := (B \cup \{k\}) \setminus \{B[h]\}$. We say that the variable x_k *enters the basis (in row h)*, and that variable $x_{B[h]}$ *exits the basis.* (Note that we should actually prove that $\tilde{B}$ is a basis; this is easy but we skip it here.)

The Simplex Method

Input: $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$, a feasible basis $B = \{B[1], \dots, B[m]\}$ per $Ax = b$, $x \geq 0$;

Output: An optimal solution $\bar{x}$ for (3.1), or we determine that the problem is unbounded.

1. Compute the tableau relative to the current basis B;
2. If $\bar{c}_j \leq 0$ for all $j \in N$, then the basic feasible solution relative to B is optimal, STOP.
3. Otherwise, choose k such that $\bar{c}_k > 0$;

   3a. If $\bar{a}_{ik} \leq 0$ $\forall i \in \{1, \dots, m\}$, then the problem is unbounded, STOP.

   3b. Otherwise, choose $h \in \{1, \dots, m\}$ such that

   $$\bar{a}_{hk} > 0 \quad \text{and} \quad \frac{\bar{b}_h}{\bar{a}_{hk}} = \min\left\{\frac{\bar{b}_i}{\bar{a}_{ik}} : i \in \{1, \dots, m\},\, \bar{a}_{ik} > 0\right\};$$

   Set $B[h] := k$, return to 1.

Case 2 $\bar{a}_{ik} \leq 0$ for every $i \in B$.

In this case, $x(t)$ is feasible for every $t \geq 0$. Since the objective value of $x(t)$ is $\bar{z} + \bar{c}_k t$, and

$$\bar{z} + \lim_{t\to+\infty} t\bar{c}_k = +\infty,$$

it follows that in this case problem (3.1) is **unbounded**.

3.1.1 Pivots

The Simplex Method described in the previous section requires to recompute, at every iteration, the tableau relative to the current basis. Here we remark that this can be easily done. Indeed, let $\tilde{B} = B \setminus \{B[h]\} \cup \{k\}$ be the new basis, where the elements of $\tilde{B}$ are $\tilde{B}[i] = B[i]$, $i \in \{1, \dots, m\} \setminus \{h\}$, $\tilde{B}[h] = k$.

The tableau $\tilde{T}$ relative to $\tilde{B}$ can be recomputed using row operations from the tableau T relative to B.

For simplicity, let us assume that $B = \{1, \dots, m\}$ and $B[i] = i$ for $i = 1, \dots, m$. Then the tableau T relative to B is of the form

$$\begin{array}{c|c|c|c|c|c|c|c|c|}
 & \multicolumn{1}{c}{} & \multicolumn{1}{c}{} & \multicolumn{1}{c}{h} & \multicolumn{1}{c}{} & \multicolumn{1}{c}{} & \multicolumn{1}{c}{k} & \multicolumn{1}{c}{} & \multicolumn{1}{c}{} \\ \cline{2-9}
 & 1 & 0 \ \dots & 0 & \dots \ 0 & \dots & -\bar{c}_k & \dots & \bar{z} \\ \cline{2-9}
 & 0 & 1 & & & & \bar{a}_{1k} & & \bar{b}_1 \\
T = & \vdots & \ddots & & & & \vdots & & \vdots \\ \cline{2-9}
 & 0 & & 1 & & & \bar{a}_{hk} & & \bar{b}_h \\ \cline{2-9}
 & \vdots & & & \ddots & & \vdots & & \vdots \\
 & 0 & & & 1 & & \bar{a}_{mk} & & \bar{b}_m \\ \cline{2-9}
\end{array}.$$

Now, the kth column of the new tableau $\tilde{T}$ must be a unit vector, whose only one is in position h. This can be accomplished by row operations as follows

$$\begin{aligned}
\mathrm{row}_h(\tilde{T}) &= \tfrac{1}{\bar{a}_{hk}}\mathrm{row}_h(T) \\
\mathrm{row}_i(\tilde{T}) &= \mathrm{row}_i(T) - \tfrac{\bar{a}_{ik}}{\bar{a}_{hk}}\mathrm{row}_h(T), \quad i = 0, \dots, m,\, i \neq h.
\end{aligned}$$

where row_0 refers to the row of the objective function.

We therefore obtain

$$\tilde{T} = \begin{array}{c} \\ \\ \end{array}\begin{array}{|c|c|c|c|c|c|c|c|}
\multicolumn{1}{c}{} & \multicolumn{1}{c}{} & \multicolumn{1}{c}{h} & \multicolumn{1}{c}{} & \multicolumn{1}{c}{} & \multicolumn{1}{c}{k} & \multicolumn{1}{c}{} & \multicolumn{1}{c}{} \\
\hline
1 & 0 \; \dots & \frac{\bar{c}_k}{\bar{a}_{hk}} & \dots \; 0 & \dots & 0 & \dots & \bar{z} + \frac{\bar{c}_k \bar{b}_h}{\bar{a}_{hk}} \\
\hline
0 & 1 & -\frac{\bar{a}_{1k}}{\bar{a}_{hk}} & & & 0 & & \bar{b}_1 - \frac{\bar{a}_{1k}\bar{b}_h}{\bar{a}_{hk}} \\
\vdots & \ddots & \vdots & & & \vdots & & \vdots \\
\hline
0 & & \frac{1}{\bar{a}_{hk}} & & & 1 & & \frac{\bar{b}_h}{\bar{a}_{hk}} \\
\hline
\vdots & & \vdots & \ddots & & \vdots & & \vdots \\
0 & & -\frac{\bar{a}_{mk}}{\bar{a}_{hk}} & 1 & & 0 & & \bar{b}_m - \frac{\bar{a}_{mk}\bar{b}_h}{\bar{a}_{hk}} \\
\hline
\end{array}$$

The operation of going from T to $\tilde{T}$ is called *pivot on the entry* (h, k) of T.

3.1.2 An example

Consider the problem

$$\begin{array}{rrcrcr}
\max z = & 130x_1 & + & 100x_2 & & \\
& 1.5x_1 & + & x_2 & \leq & 27 \\
& x_1 & + & x_2 & \leq & 21 \\
& 0.3x_1 & + & 0.5x_2 & \leq & 9 \\
& x_1 & , & x_2 & \geq & 0
\end{array}$$

We first write the above problem in standard equality form by adding slack variables x_3, x_4, x_5.

$$\begin{array}{rrcrcrcrcrcr}
\max z = & 130x_1 & + & 100x_2 & & & & & & & & \\
& 1.5x_1 & + & x_2 & + & x_3 & & & & & = & 27 \\
& x_1 & + & x_2 & & & + & x_4 & & & = & 21 \\
& 0.3x_1 & + & 0.5x_2 & & & & & + & x_5 & = & 9 \\
& x_1 & , & x_2 & , & x_3 & , & x_4 & , & x_5 & \geq & 0
\end{array}$$

Note that the basis $\{3, 4, 5\}$ is feasible. The corresponding basic solution is $x_1 = 0$, $x_2 = 0$, $x_3 = 27$, $x_4 = 21$, $x_5 = 9$, with objective value 0. This solution correspond to the vertex $(0, 0)$ in the original problem. The problem in tableau form relative to the basis $B := \{B[1] := 3, B[2] := 4, B[3] := 5\}$ is

$$\begin{array}{rcrcrcrcrcrcr}
z & - & 130x_1 & - & 100x_2 & & & & & & & = & 0 \\
& & 1.5x_1 & + & x_2 & + & x_3 & & & & & = & 27 \\
& & x_1 & + & x_2 & & & + & x_4 & & & = & 21 \\
& & 0.3x_1 & + & 0.5x_2 & & & & & + & x_5 & = & 9
\end{array}$$

Every iteration of the Simplex Method will comprise three steps: choosing the entering variable, choosing the exiting variable, pivoting.

Iteration 1

Step 1: Choose the entering variable. *Choose x_k with positive reduced cost.*

In the previous example, the reduced costs of x_1 and x_2 are 130 and 100, respectively. They both could be chosen. Let us choose x_1.

Step 2: Choose the exiting variable. *If x_k enters the basis, choose as exiting variable one that is currently in basis in a row that gives the minimum quotient*

$$\min_{i : \bar{a}_{ij} > 0} \left\{ \frac{\bar{b}_i}{\bar{a}_{ij}} \right\}.$$

In the example, $\min\{\frac{27}{1.5}, \frac{21}{1}, \frac{9}{0.3}\} = 18$, which is attained in the row whose basic variable is x_3. Thus we select x_3 as the exiting variable. Then new basis is $\{1.4, 5\}$.

Recall that the minimum quotient rule is justified by the following argument. We intend to increase x_1 as much as possible, while keeping unchanged to 0 the value of the remaining nonbasic variables (in this case, only x_2) without violating the constraints. If we increase x_1 by t, in order to satisfy the equality constraints the basic variables must be modified as follows:

$$\begin{aligned} x_3 &= 27 - 1.5t \\ x_4 &= 21 - t \\ x_5 &= 9 - 0.3t \end{aligned}$$

Since the above values must be nonnegative, the maximum value of t that we can choose is $\min\{\frac{27}{1.5}, \frac{21}{1}, \frac{9}{0.3}\}$.

Step 3: Pivot We need to bring the problem in tableau form with respect to the new basis $\{1.4, 5\}$. At this stage, the row of the tableau containing x_3 is $1.5x_1 + x_2 + x_3 = 27$, and x_1 must appear in that row with a 1 coefficient. Thus, we divide the equation by 1.5, obtaining the new tableau row $x_1 + \frac{2}{3}x_2 + \frac{2}{3}x_3 = 18$. The second row of the tableau is $x_1 + x_2 + x_4 = 21$. In the new tableau x_1 must have a 0 coefficient in the second row, thus, subtracting the first row of the new tableau to the second row of the old tableau we obtain the equation $\frac{1}{3}x_2 - \frac{2}{3}x_3 + x_4 = 3$. Proceeding similarly for the 0th and 3rd row of the tableau, we obtain the problem in tableau form with respect to the new basis.

$$\begin{array}{rrrrrrrrrrrr} z & & & - & \frac{40}{3}x_2 & + & \frac{260}{3}x_3 & & & & = & 2340 \\ & x_1 & + & & \frac{2}{3}x_2 & + & \frac{2}{3}x_3 & & & & = & 18 \\ & & & & \frac{1}{3}x_2 & - & \frac{2}{3}x_3 & + & x_4 & & = & 3 \\ & & & & 0.3x_2 & - & 0.2x_3 & & & + \; x_5 & = & 3.6 \end{array}$$

The new basic solution is $(18, 0, 0, 3, 6)$, with value 2340.

Iteration 2

In the above tableau, the only variable with positive reduced cost is x_2, with reduced cost $\frac{40}{3}$, hence we have to choose x_2 has the entering variable. To choose the exiting variable, we apply the minimum quotient rule: $\min\{\frac{18}{2/3}, \frac{3}{1/3}, \frac{3.6}{0.3}\} = 9$. The minimum is achieved in the row corresponding to the basic variable x_4, thus x_4 exits the basis. The new basis is $\{1, 2, 5\}$.

After we perform a pivot, we obtain the new tableau:

$$\begin{array}{rrrrrrrrrrr} z & & & + & 60x_3 & + & 40x_4 & & & = & 2460 \\ & x_1 & & + & 2x_3 & - & 2x_4 & & & = & 12 \\ & & x_2 & - & 2x_3 & + & 3x_4 & & & = & 9 \\ & & & - & 0.4x_3 & - & 0.9x_4 & + & x_5 & = & 0.9 \end{array}$$

Iteration 3

At this point, there are no variables with positive reduced cost. We stop with the optimal solution $x_1 = 12$, $x_2 = 9$, $x_3 = x_4 = 0$, $x_5 = 0.9$. The optimal value of the problem is 2460.

Note that the Simplex Method, in the space of the original variables x_1, x_2, followed the sequence of vertices $(0, 0)$, $(18, 0)$, $(12, 9)$.

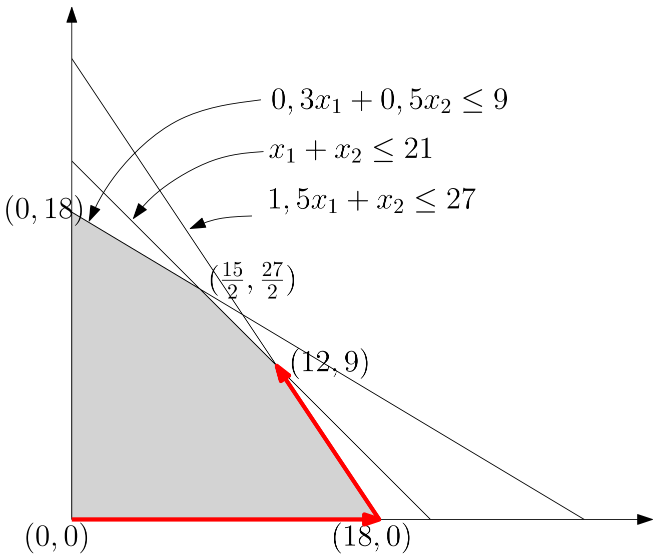


Figure 3.1: Representation of the feasible region of the problem in the example, and path followed by the Simplex Method.

Example of an unbounded problem

$$\begin{array}{cccccccc} z & & & + & 1.5x_3 & - & 0.25x_4 & = & 3 \\ & x_1 & & + & 0.5x_3 & - & 0.25x_4 & = & 2 \\ & & x_2 & - & 0.5x_3 & - & 0.25x_4 & = & 1 \end{array}$$

The current basic solution is $(2,1,0,0)$ The only variable eligible to enter the basis is x_4, because it is the only one with positive reduced cost. If we increase the value of x_4 to t leaving unchanged to 0 the value of the remaining nonbasic variables, then in order to satisfy the equality constraints we need to modify the values of the basic variables as follows.

$$\begin{aligned} x_1(t) &= 2 + 0.25t \\ x_2(t) &= 1 + 0.25t \end{aligned}$$

Note that $x_1(t)$ and $x_2(t)$ are nonnegative *for any* $t \geq 0$. We can therefore increase x_4 arbitrarily. As t increases, also the objective value of the solution increases. This implies that the problem is unbounded.

Our discussion of the Simplex Method leaves open two fundamental questions.

1. Does the Simplex Method terminate?

2. How does one find a basic feasible solution to start?

These will be addressed in the next two sections.

3.2 Degenerate bases and termination

The Simplex Method terminates when one of the conditions at point 2 or point 3a occurs. It is clear that, when the method terminates, the output is correct. However, we need to argue that the method always terminate.

First of all, note that, if at every iteration the value

$$\bar{t} := \min_{i \in \{1,\ldots,m\} : \bar{a}_{ik} > 0} \left\{ \frac{\bar{b}_i}{\bar{a}_{ik}} \right\} \tag{3.9}$$

was positive, then the objective value of the current basic feasible solution would strictly increase at every iteration. Now, since the basic feasible solution at any iteration is completely determined by the current basis, this means that the same basis is never visited twice. But the number of bases is finite, hence the method would terminate after a finite numbers of iterations in this case.

Thus problems may arise only in the case that, at some iteration, the objective function value does not change. Note that $\bar{t} = 0$ if and only if there exists some index $i \in \{1, \dots, m\}$ such that $\bar{b}_i = 0$ and $\bar{a}_{ik} > 0$. Note that, when this happens, $\bar{x}_{B[i]} = b_i = 0$. Also note that the new basis $B \setminus \{B[h]\} \cup \{k\}$ determined at point 3b of the Simplex Method defines the same basic feasible solution as B. In other words, we have changed the current basis, but the current solution remains the same. In this case we say that we have performed a *degenerate iteration.*

Definition 3.3. *A feasible basis B is said to be* non-degenerate *if the basic feasible solution $\bar{x}$ relative to B satisfies $\bar{x}_i > 0$ for every $i \in B$, that is, if and only if $\bar{b} = A_B^{-1} b > 0$. Otherwise, we say that the basis is* degenerate.

So, as we said earlier, if it happened that, during the execution of the Simplex Method, all bases visited are non-degenerate, then the objective value of the current solution increases strictly at every iteration and the method terminates after a finite number of iterations.

However, this does not happen in general. Indeed, it could happen that the Simplex Method "gets stuck" in an infinite loop, cyclically visiting the same sequence of bases indefinitely. This phenomenon is known as *cycling.*

How can we prevent this from happening? We should realize at this point that the description we gave of the Simplex Method does not define a precise algorithm, because we have not established any rule about how to choose the entering and exiting variable. In general, there might be several variables with positive reduced costs, and we have to choose one of them as entering variable. Analogously, once we have chosen the entering variable x_k, there might be several variables that are candidates to exit the basis, namely all the variables x_h satisfying $\bar{a}_{hk} > 0$ and $\frac{\bar{b}_h}{\bar{a}_{hk}} = \bar{t}$. A rule that establishes a unique choice for the entering and exiting variables is called a *pivot rules.* While not every pivot rule one could conceive will guarantee termination, there are several "anti-cycling" rules.

One of the most known such rules is the so called *Bland's rule* (named after its inventor, Bob Bland): *among all variables with positive reduced cost, choose as entering variable the one with smallest possible index; among all variables that minimize the quotient in (3.9) choose as exiting variable the one with smallest possible index.* Formally, Bland's rule is defined as follows.

Bland's rule: *At any iteration of the Simplex Method, relative to a feasible basis $B = \{B[1], \dots, B[m]\}$:*

- *Among all variables with positive reduced cost, choose as entering variable the variable x_k such that the index k is the smallest possible.*
- *Let $\bar{t} = \min\{\frac{\bar{b}_i}{\bar{a}_{ik}} \,|\, \bar{a}_{ik} > 0\}$. Choose as exiting variable the variable $x_{B[h]}$ such that $\bar{a}_{hk} > 0$, $\frac{\bar{b}_h}{\bar{a}_{hk}} = \bar{t}$, and such that $B[h]$ is smallest possible.*

Example 3.4. Consider the following LP problem in tableau form with respect to the basis $\{1, 5, 4\}$ (thus $B[1] = 1$, $B[2] = 5$, $B[3] = 4$).

$$\begin{array}{ccccccccccc}
z & & & - & 3x_2 & - & 7x_3 & & & & = & 26 \\
& x_1 & + & & \frac{3}{2}x_2 & + & \frac{2}{3}x_3 & & & & = & 18 \\
& & & & 0.4x_2 & - & 0.2x_3 & & & + \; x_5 & = & 3.6 \\
& & & & \frac{1}{3}x_2 & - & \frac{2}{3}x_3 & + & x_4 & & = & 3
\end{array}$$

The variables with positive reduced cost are x_2 and x_3, with reduced cost 3 and 7, respectively. Applying bland rule, we have to choose as entering variable the one with smallest index among the two, which is x_2.

To choose the exiting variable, we need to apply the minimum quotient rule. Note that $\bar{t} = \min\{\frac{18}{3/2}, \frac{3.6}{0.4}, \frac{3}{1/3}\} = 9$. Such value is attained in two tableau rows: the second (relative to the variable

x_5) and the third (relative to the variable x_4). Hence we select as exiting variable the one among x_4 and x_5 that has smallest index, which is x_4.

Theorem 3.5 (Termination of the Simplex Method with Bland's rule)**.** *The Simplex Method with Bland's pivot rule terminates for every possible problem and every possible choice of starting feasible basis.*

We do not prove this theorem here.

3.3 The two-phase Simplex Method

The Simplex Method requires a feasible basis to start. Suppose we want to find a feasible basis to the problem

$$\begin{aligned} \max c^\top x & \\ Ax = & \ b \\ x \geq & \ 0 \end{aligned} \tag{3.10}$$

where A has m rows and n columns. Note first that we may assume that the right-hand-sides are all nonnegative (i.e. $b \geq 0$), since we can just change the signs of all coefficients in the constraints for which b_i is negative.

To check if the problem has a feasible solution, we construct an *auxiliary problem* as follows. The auxiliary problem will have $n + m$ variables, namely: the original variables $x_1, \ldots, x_n$, and the m *auxiliary variables* $x_{n+1}, \ldots, x_{n+m}$, one for each constraint. If we denote by x_A the vector $(x_{n+1}, \ldots, x_{n+m})^\top$ of auxiliary variables, the auxiliary problem is the following.

$$\begin{array}{rrcl} \max & -\sum_{i=1}^m x_{n+i} & & \\ & Ax + Ix_A & = & b \\ & x, x_A & \geq & 0 \end{array} \tag{3.11}$$

The auxiliary problem has the following properties.

(a) It is always feasible, since $\bar{x} := 0$ $\bar{x}_A := b$ is a feasible solution for the auxiliary problem (this is true because of our assumption that $b \geq 0$). In fact, the above solution is a basic solution relative to the basis $\tilde{B} := \{n + 1, \ldots, n + m\}$.

(b) The objective function is the opposite of the sum of all auxiliary variables. Since the auxiliary variables are nonnegative, minus their sum is always nonpositive. In particular, the objective function takes nonpositive value on every feasible solution. Hence the optimal value of the problem is always bounded above by 0.

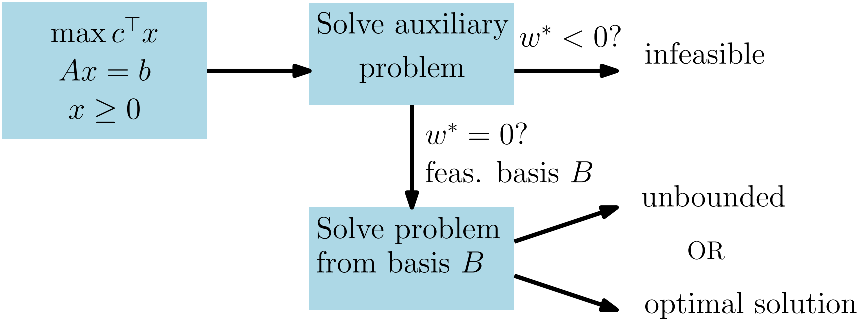
(c) A vector $\bar{x} \in \mathbb{R}^n$ is a feasible solution for (3.10) if and only if $\binom{\bar{x}}{0}$ is feasible for (3.11).

The so called **Phase 1** consists in solving (3.11) by applying the Simplex Method starting from the feasible basis $\tilde{B}$. Since the problem bounded, upon termination we obtain an optimal basic solution for the auxiliary problem, say $\binom{x^*}{x_A^*}$, relative to some basis $B^* \subseteq \{1, \ldots, n + m\}$. As mentioned above, the optimal value $w^* = -\sum_{i=1}^m x_{n+i}^*$ is always nonnegative.

There are two cases.

- $w^* = 0$: in this case it must be $x_A^* = 0$, therefore x^* is feasible for (3.10).
- $w^* < 0$: In this case (3.1) is infeasible, because if $\bar{x}$ was a feasible solution for (3.1), then $\binom{\bar{x}}{0}$ would be feasible for (3.11) with value $0 > w^*$, contradicting the fact that $\binom{x^*}{x_A^*}$ was optimal.

We note that, when $w^* = 0$, the vector x^* is indeed a basic feasible solution. If this is the case, we have determined a feasible basis to start the Simplex Method.



Two phase method

Input: $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$;

Output: One of the following

- An optimal solution for (3.10),
- The problem is infeasible,
- The problem is unbounded.

1. Construct the auxiliary problem (3.11) and solve it with the Simplex Method starting from the basis $\{n+1, \ldots, n+m\}$. **(Phase 1)**
2. If $w^* < 0$, (3.10) is **infeasible**. STOP.
3. If $w^* = 0$, then we have determined a feasible basis B for (3.10).
4. Apply the Simplex Method to the original problem (3.10) starting from the basis B computed in 3. **(Phase 2)**
5. Return either the **optimal solution** for (3.10), or the fact that (3.10) is **unbounded**, STOP.

Note that, since the Simplex Method terminates (if we apply an anti-cycling pivoting rule, such as for example Bland's rule), it follows from the two-phases method that there are only three possible outcomes for an LP problem. We thus obtain the Fundamental Theorem of Linear Programming, as already stated in MA423.

Theorem 3.6 (Fundamental Theorem of Linear Programming)**.** *For any linear programming problem, exactly one of the following holds.*

1. *The problem has a finite optimum;*
2. *The problem is infeasible;*
3. *The problem is unbounded.*

Furthermore, whenever the Simplex Method is applied to an LP problem that has a finite optimum, it outputs an optimal solution that is basic. This proves Theorem 2.16.

3.4 Simplex and duality

3.4.1 Dual basic solutions and strong duality

Let us again consider the problem (P) and its dual (D):

$$\begin{array}{rl} \max c^\top x & \\ Ax = & b \quad (P) \\ x \geq & 0 \end{array} \qquad\qquad \begin{array}{rl} \min b^\top y & \\ A^\top y \geq & c. \quad (D) \end{array}$$

Let B be a basis (possibly infeasible) of $Ax = b$, $x \geq 0$. The basic solution determined by B is

$$\bar{x} = \begin{pmatrix} \bar{x}_B \\ \bar{x}_N \end{pmatrix} = \begin{pmatrix} A_B^{-1} b \\ 0 \end{pmatrix}.$$

Let us define the vector

$$\bar{y} = (A_B^{-1})^\top c_B.$$

Note that

$$c^\top \bar{x} = c_B^\top \bar{x}_B = c_B^\top A_B^{-1} b = b^\top \bar{y}.$$

Hence, if $\bar{x}$ was feasible for (P) and $\bar{y}$ was feasible for (D), then by the weak duality theorem we would argue that $\bar{x}$ is optimal for (P) and $\bar{y}$ is optimal for (D). Now:

$\bar{x}$ is (primal) feasible if and only if it is nonnegative, that is, if and only if $\bar{b} = A_B^{-1} b \geq 0$.

$\bar{y}$ is (dual) feasible if and only it satisfies $A^\top \bar{y} \geq c$. Splitting the latter constraints between the ones relative to indices in B and the ones relative to indices in N, we can write

$$A_B^\top \bar{y} \;\geq\; c_B \tag{3.12}$$
$$A_N^\top \bar{y} \;\geq\; c_N \tag{3.13}$$

However, by definition of $\bar{y}$ we have that $A_B^\top \bar{y} = A_B^\top (A_B^{-1})^\top c_B = c_B$, hence condition (3.12) is always satisfied. Thus $\bar{y}$ is feasible if and only if $A_N^\top (A_B^{-1})^\top c_B \geq c_N$.

Rearranging the terms, we obtain that $\bar{y}$ is (dual) feasible if and only if

$$c_N - A_N^\top (A_B^{-1})^\top c_B \leq 0.$$

Note that $c_N - A_N^\top (A_B^{-1})^\top c_B = \bar{c}_N$ is the vector of reduced costs relative to B defined in (3.4).

It follows that $\bar{y}$ is dual feasible if and only if the reduced costs associated to the basis B are nonpositive, i.e.

$$\bar{c}_N \leq 0.$$

This is precisely the condition verified when the Simplex Method terminates with an optimal solution. Thus, when the Simplex Method terminates with an optimal basic solution $\bar{x}$, and a basis B that determines such solution, the vector $\bar{y} = (A_B^{-1})^\top c_B$ is a dual optimal solution, which certifies the optimality of $\bar{x}$.

Definition 3.7. *Let B be a basis for the LP problem (P).*
We say that B is primal feasible *if the corresponding basic solution is feasible, i.e. if $A_B^{-1} b \geq 0$.*
We say that B is dual feasible *if the corresponding dual solution is feasible, i.e. if $c_N - A_N^\top (A_B^{-1})^\top c_B \leq 0$.*
If B is both primal feasible and dual feasible, then we say that B is an optimal basis.

Therefore the Simplex Method maintains, at every iteration, a basis that is primal feasible (but not dual feasible), and terminates when the basis becomes also dual feasible, and therefore optimal (unless the problem is unbounded). Hence, if (P) has a finite optimum, the Simplex Method will determine an optimal basis, and so it will determine a primal feasible solution and a dual feasible solution with the same value.

Since the Simplex Method with Bland's rule terminates, this implies the Strong Duality Theorem (Theorem 1.7). We state it again for convenience.

Theorem 3.8 (Strong Duality Theorem)**.** *If (P) has an optimal solution x^*, then its dual (D) has an optimal solution y^*, and $c^\top x^* = b^\top y^*$.*

Example. Consider the example of Section 3.1.2.

$$A = \begin{bmatrix} 1.5 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 0.3 & 0.5 & 0 & 0 & 1 \end{bmatrix}, \qquad b = \begin{bmatrix} 27 \\ 21 \\ 9 \end{bmatrix}, \qquad c = \begin{bmatrix} 130 \\ 100 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

The basis we find after the first iteration is $B = \{1, 4, 5\}$. We have

$$A_B^{-1} = \begin{bmatrix} \frac{2}{3} & 0 & 0 \\ -\frac{2}{3} & 1 & 0 \\ -\frac{1}{5} & 0 & 0 \end{bmatrix}$$

The basis B is primal feasible, and in fact

$$A_B^{-1} b = \begin{bmatrix} 18 \\ 3 \\ 3.6 \end{bmatrix} \geq \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

which correspond to the basic feasible solution $(18, 0, 0, 3, 3.6)^\top$

The dual solution associated to B is

$$\bar{y} = (A_B^{-1})^\top c_B = \begin{bmatrix} \frac{2}{3} & -\frac{2}{3} & -\frac{1}{5} \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 130 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{260}{3} \\ 0 \\ 0 \end{bmatrix}$$

Such solution is not feasible in the dual. Indeed, it violates the second dual constraint

$$y_1 + y_2 + 0.5 y_3 \geq 100$$

since

$$1 \cdot \frac{260}{3} + y_2 \cdot 0 + 0.5 \cdot 0 = \frac{260}{3} < 100.$$

Note that the amount by which $\bar{y}$ violates the constraint is $\frac{40}{3}$, which is precisely the reduced costs of x_2 with respect to the basis B.

At the final iteration, the basis computed in example is $B = \{1, 2, 5\}$. We have

$$A_B^{-1} = \begin{bmatrix} 2 & -2 & 0 \\ -2 & 3 & 0 \\ 0.4 & -0.9 & 1 \end{bmatrix}$$

The basis B is primal feasible, and in fact

$$A_B^{-1} b = \begin{bmatrix} 12 \\ 9 \\ 0.9 \end{bmatrix} \geq \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

which correspond to the basic feasible solution $(12, 9, 0, 0, 0.9)^\top$

The dual solution associated to B is

$$\bar{y} = (A_B^{-1})^\top c_B = \begin{bmatrix} 2 & -2 & 0.4 \\ -2 & 3 & -0.9 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 130 \\ 100 \\ 0 \end{bmatrix} = \begin{bmatrix} 60 \\ 40 \\ 0 \end{bmatrix}.$$

It can be verified that $\bar{y}$ is feasible in the dual. Thus $B = \{1, 2, 5\}$ is an optimal basis.

Observe that the value of $(12, 9, 0, 0, 0.9)^\top$ in the primal is $130 \cdot 12 + 100 \cdot 9 = 2460$, and the value of $\bar{y}$ in the dual is $27 \cdot 60 + 21 \cdot 40 = 2460$.

3.4.2 The Dual Simplex Method

The Dual Simplex Method keeps at every iteration a dual feasible basis, and terminates when it determines a basis that is also primal feasible (and thus optimal). Hence the Dual Simplex can be interpreted as the Simplex Method applied to the dual rather than on the primal problem.

As usual, let us consider the primal/dual pair of linear programming problems

$$\begin{array}{rl} \max c^\top x & \\ Ax = & b \\ x \geq & 0 \end{array} \quad (P) \qquad \begin{array}{rl} \min b^\top y & \\ A^\top y \geq & c. \end{array} \quad (D)$$

Let $B = \{B[1], \dots, B[m]\}$ be a dual feasible basis (but possibly not primal feasible). Let the problem in tableau for with respect to such basis be

$$\begin{array}{llrcll} z^* = \max & z & & & & \\ & z & -\sum_{j\in N} \bar{c}_j x_j & = & \bar{z} & \\ & x_{B[i]} & +\sum_{j\in N} \bar{a}_{ij} x_j & = & \bar{b}_i, & i = 1, \dots, m \\ & x & & \geq & 0. & \end{array} \tag{3.14}$$

Since B is dual feasible the reduced costs with respect to B are all nonpositive, so $\bar{c}_j \leq 0$ for every $j \in N$. Note that $\bar{z} = c_B^\top A_B^{-1} b$ is the value of the dual solution $\bar{y} = (A_B^\top)^{-1} c_B$ associated with B. Since $\bar{y}$ is a dual feasible solution and the dual problem is a minimization problem, it follows that $z^* \leq \bar{z}$

The basic primal solution $\bar{x}$ relative to B is defined, as usual, by

$$\begin{array}{rcl} \bar{x}_{B[i]} & = & \bar{b}_i; \quad i = 1, \dots, m \\ \bar{x}_j & = & 0; \quad j \in N. \end{array}$$

If $\bar{b}_i \geq 0$ for $i = 1, \dots, m$, then $\bar{x}$ is feasible, and thus B is a primal feasible basis, and thus an optimal basis. In this case, $\bar{x}$ is an optimal solution and we are done.

Suppose then that there exists some index $h \in \{1, \dots, m\}$ such that $\bar{b}_h < 0$. We select the corresponding basic variable $x_{B[h]}$ to leave the basis. We need to select an entering variable in such a way that the new basis will remain dual feasible. There are two cases.

Case 1: $\bar{a}_{hj} \geq 0$ for every $j \in N$.

In this case, the coefficients of the variables in the tableaux equation

$$x_{B[h]} + \sum_{j\in N} \bar{a}_{hj} x_j = \bar{b}_h$$

are all nonnegative, while the right-hand-side is negative. It follows that no vector $x \geq 0$ can satisfy such equation, since we should have $x_{B[h]} + \sum_{j\in N} \bar{a}_{hj} x_j \geq 0$, which cannot be the case since $b_h < 0$. In this case, we conclude that the primal problem is infeasible.

Case 2: $\bar{a}_{hj} < 0$ for some $j \in N$.

We need to determine an index k such that the new basis $\tilde{B} = B \cup \{k\} \setminus \{B[h]\}$ is again dual feasible. This is the case if the vector $\tilde{c}$ of the reduced costs relative to the new basis $\tilde{B}$ is nonpositive. Let $\tilde{N} = \{1, \dots, n\} \setminus \tilde{B}$. As one can verify by performing a pivot on the entry (h, k) of the tableau relative to B, the vector $\tilde{c}_{\tilde{N}}$ is given by

$$\begin{array}{rcl} \tilde{c}_{B[h]} & = & -\dfrac{\bar{c}_k}{\bar{a}_{hk}}, \\ \tilde{c}_j & = & \bar{c}_j - \dfrac{\bar{c}_k}{\bar{a}_{hk}} \bar{a}_{hj}, j \in N \setminus \{k\} \end{array}$$

Since $\bar{c}_k \leq 0$, the reduced cost $\tilde{c}_{B[h]}$ is less than or equal to zero if and only if $\bar{a}_{hk} < 0$. We must therefore choose an index k such that $\bar{a}_{hk} < 0$. Furthermore, we need to satisfy

$$\tilde{c}_j = \bar{c}_j - \frac{\bar{c}_k}{\bar{a}_{hk}} \bar{a}_{hj} \leq 0, \qquad j \in N \setminus \{k\}.$$

For every $j \in N \setminus \{k\}$ such that $\bar{a}_{hj} \geq 0$, the above condition is always verified, since $\frac{\bar{c}_k}{\bar{a}_{hk}}\bar{a}_{hj} \leq 0$ and thus $\tilde{c}_j \leq \bar{c}_j \leq 0$. If $\bar{a}_{hj} < 0$, then $\tilde{c}_j \leq 0$ if and only if

$$\frac{\bar{c}_k}{\bar{a}_{hk}} \leq \frac{\bar{c}_j}{\bar{a}_{hj}}.$$

Thus we need to choose k such that

$$\frac{\bar{c}_k}{\bar{a}_{hk}} \leq \frac{\bar{c}_j}{\bar{a}_{hj}} \text{ for every } j \in N \setminus \{k\} \text{ such that } \bar{a}_{hj} < 0.$$

This means that we need to choose $k \in N$ so that $a_{hk} < 0$ and

$$\frac{\bar{c}_k}{\bar{a}_{hk}} = \min\left\{ \frac{\bar{c}_j}{\bar{a}_{hj}} \ : \ j \in N, \, \bar{a}_{hj} < 0 \right\}.$$

For such a choice of k, the new basis $\tilde{B}$ is dual feasible. Note that the objective value of the dual solution relative to the new basis $\tilde{B}$ is

$$c_{\tilde{B}} A_{\tilde{B}}^{-1} b = \bar{z} + \frac{\bar{c}_k}{\bar{a}_{hk}} \bar{b}_h \leq \bar{z},$$

where the inequality holds because $\bar{c}_k \leq 0$, $\bar{a}_{hk} < 0$ and $\bar{b}_h < 0$.

In other words, the value of the new dual solution is less than or equal to the value of the previous one, which means that we are finding better and better dual solutions at every iterations (since the dual is a minimization problem). The method will terminate when the basis is primal feasible. We present the Dual Simplex Method in the following table.

Dual Simplex Method

Input: $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$, a dual feasible basis $B = \{B[1], \dots, B[m]\}$;

Output: Either an optimal solution $\bar{x}$ for (P), or we determine that (P) is infeasible.

1. Compute the tableau with respect to the current basis B;
2. If $\bar{b} \geq 0$, then B is optimal, STOP.
3. Otherwise, choose an index h such that $\bar{b}_h < 0$;

   3a. If $\bar{a}_{hj} \geq 0 \ \forall j \in N$, then the problem is infeasible, STOP.

   3b. Otherwise choose $k \in N$ such that

   $$\bar{a}_{hk} < 0 \quad \text{ and } \quad \frac{\bar{c}_k}{\bar{a}_{hk}} = \min\left\{ \frac{\bar{c}_j}{\bar{a}_{hj}} \ : \ j \in N, \, \bar{a}_{hj} < 0 \right\};$$

   Set $B[h] := k$, return to 1.

Note that the for the Dual Simplex Method we are presented with the usual problem of termination: if we are not careful with our choice of pivoting rule, the method might not terminate. However the method is guaranteed to terminated if we apply Bland's dual rule, where we choose as exiting variable the one of minimum index among the candidates to leave the basis, and we choose as entering variable the one of minimum index among the candidates to enter the basis.

Bland's dual rule: *At any iteration of the Dual Simplex Method, relative to a dual feasible basis* $B = \{B[1], \dots, B[m]\}$*:*

- *Choose as exiting variable the variable* $x_{B[h]}$ *such that* $\bar{b}_h < 0$ *and* $B[h]$ *is minimum.*
- *Let* $\bar{t} = \min\{\frac{\bar{c}_j}{\bar{a}_{hj}} : j \in N, \bar{a}_{hj} < 0\}$. *Choose as entering variable the variable* x_k *such that* $\bar{a}_{hk} < 0$, $\frac{\bar{c}_k}{\bar{a}_{hk}} = \bar{t}$, *and such that* k *is smallest possible.*