

Linear Programming in Combinatorics

Amritanshu Prasad

THE INSTITUTE OF MATHEMATICAL SCIENCES, CHENNAI.

HOMI BHABHA NATIONAL INSTITUTE, MUMBAI.

Email address: `amri@imsc.res.in`

CHAPTER 1

Introduction to Linear Programming

An excellent reference for the material in this section is the book of Gärtner and Matousek [1].

1. Feasibility and Optimization

A *linear program in equational form* consists of a set of variables $\mathbf{x} = (x_1, \dots, x_n)$, an $m \times n$ matrix A with real entries, a bound vector $\mathbf{b} = (b_1, \dots, b_m)$, and an objective vector $\mathbf{c} = (c_1, \dots, c_n)$. The *linear program* is the problem:

$$(LP) \quad \text{maximize } \mathbf{c}^T \mathbf{x} \text{ subject to } \mathbf{x} \geq 0 \text{ and } A\mathbf{x} = \mathbf{b}.$$

The set

$$P(A, \mathbf{b}) = \{\mathbf{x} \in \mathbf{R}^n \mid \mathbf{x} \geq 0, A\mathbf{x} = \mathbf{b}\}$$

is called the *polytope* of all *feasible solutions* to (LP). The function $\mathbf{x} \mapsto \mathbf{c}^T \mathbf{x}$ is called the *objective function*. An *optimal solution* is a vector $\mathbf{x}_0 \in P(A, \mathbf{b})$ such that $\mathbf{c}^T \mathbf{x} \leq \mathbf{c}^T \mathbf{x}_0$ for every $\mathbf{x} \in P(A, \mathbf{b})$. Sometimes we will only be interested in the set $P(A, \mathbf{b})$ of feasible solutions, which does not depend on the objective vector.

Assume that \mathbf{b} lies in the column space of A (for otherwise, we would have $P(A, \mathbf{b}) = \emptyset$), and that the rows of A are linearly independent (if not, we could discard redundant rows to achieve this). For a subset $B \subset [n]$, let A_B denote the submatrix of A consisting of columns from B . We say that B is a *basic set* if B has m elements and A_B has rank m . For $\mathbf{x} \in \mathbf{R}^n$ define:

$$\text{supp}(\mathbf{x}) = \{1 \leq j \leq n \mid x_j \neq 0\}.$$

DEFINITION 1.1 (Basic feasible solution). A *basic feasible solution* to (LP) is a feasible solution $\mathbf{x} \in P(A, \mathbf{b})$ such that $\text{supp}(\mathbf{x})$ is contained in a basic set.

Clearly, a feasible solution \mathbf{x} is basic if and only if the submatrix $A_{\text{supp}(\mathbf{x})}$ has linearly independent columns.

EXAMPLE 1.2 (Transportation polytopes). Take $n = rs$, indexing the rs variables as $\mathbf{x} = (x_{ij})_{1 \leq i \leq r, 1 \leq j \leq s}$, a rectangular array with r rows and s columns.

Let $\mathbf{a} \in \mathbf{R}_{\geq 0}^r$, and $\mathbf{b} \in \mathbf{R}_{\geq 0}^s$ be fixed vectors. As constraints, say that the row sums and column sums of \mathbf{x} are specified by the vectors \mathbf{a} and \mathbf{b} .

$$(R) \quad \sum_j x_{ij} = a_i \text{ for } i = 1, \dots, r,$$

$$(C) \quad \sum_i x_{ij} = b_j \text{ for } j = 1, \dots, s.$$

These $r + s$ constraints are not independent—the left hand side of the sum of the row sum constraints (R) is equal to the left hand side of the sum of the column sum constraints (C), being a constraint on the sum of all entries of the matrix.

If $\sum a_i \neq \sum b_j$, then there are no feasible solutions of this program. Thus we assume that the vectors \mathbf{a} and \mathbf{b} have equal sums. Now, removing, for instance, the last column constraint gives an $(r + s - 1) \times rs$ matrix A of rank $r + s - 1$, which is equivalent to the original system. We will denote the resulting polytope of feasible solutions by $\mathbf{M}_{\mathbf{ab}}$. This polytope is known as the *transportation polytope with margins* (\mathbf{a}, \mathbf{b}) . For an excellent survey on transportation polytopes, see [2].

When $r = s = 2$, the constraints can be expressed in matrix form as

$$(1) \quad \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_{11} \\ x_{12} \\ x_{21} \\ x_{22} \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ b_1 \end{pmatrix}.$$

Every submatrix of A with three columns is non-singular. Thus there are four possible basic sets.

EXAMPLE 1.3 (Birkhoff polytopes). When $r = s$ and $\mathbf{a} = \mathbf{b} = (1, \dots, 1)$, then the resulting polytope is known as the r th *Birkhoff polytope*. Consider the case where $r = 2$, so the constraints are given by (1) with $a_1 = a_2 = b_1 = 1$. It is easy to see that there are only two basic solutions, given by the permutation matrices:

$$\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \text{ or } \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Each basic feasible solution corresponds to a two possible basic subsets.

EXERCISE 1.4. Let B be a basic subset of $[r] \times [s]$ for the transportation polytope. By definition, B has cardinality $r + s - 1$. Given B , define a bipartite graph Γ_B on the set $\{1, \dots, r\} \sqcup \{1', \dots, s'\}$ by joining i to j' if $(i, j) \in B$. Show that Γ_B is a spanning tree for the complete bipartite graph $K_{r,s}$. Show that this construction gives rise to a bijection from the set of basic subsets of $[r] \times [s]$ onto the set of spanning trees of $K_{r,s}$.

EXERCISE 1.5. Consider the r th Birkhoff polytope, $\mathbf{M}_{(1^r), (1^r)}$. Let $\sigma \in S_r$ be a permutation and suppose \mathbf{x} is the permutation matrix $x_{ij} = \delta_{i\sigma(i)}$. Then $\text{supp}(\mathbf{x}) = \{(i, \sigma_i) \mid i \in [r]\}$. The first r rows of the corresponding column vectors of A are just the coordinate vectors of \mathbf{R}^r . Therefore each permutation matrix is a basic feasible solution. Which spanning trees does it correspond to?

LEMMA 1.6. For each basic subset $B \subset [n]$ of (LP), there exists at most one basic feasible solution \mathbf{x} with $\text{supp}(\mathbf{x}) \subset B$.

In the situation described in the preceding lemma, we say that \mathbf{x} is the *basic solution corresponding to B*.

PROOF. Let \mathbf{x}_B denote the vector $(x_i)_{i \in B}$. The matrix A_B is non-singular, so the equation $A_B \mathbf{x}_B = \mathbf{b}$ has at most one solution. Solutions \mathbf{x} of $A\mathbf{x} = \mathbf{b}$ with $\text{supp}(\mathbf{x}) \subset B$ are in bijection with solutions of $A_B \mathbf{x}_B = \mathbf{b}$ (set $x_j = 0$ for $j \notin B$ to get \mathbf{x} from \mathbf{x}_B). Therefore $A\mathbf{x} = \mathbf{b}$ also has at most one solution. \square

REMARK 1.7. The same basic feasible solution could be obtained from different basic sets. For example, each basic solution for $r = 2$ in Example 1.3 corresponds to two basic sets. Also not every basic set B admits a basic feasible solution. For example, for the 3rd Birkhoff polytope, $B = \{(1, 1), (1, 2), (1, 3), (2, 1), (3, 1)\}$ is a basic set with no feasible solution.

THEOREM 1.8 (Existence of basic optimal solutions). *For a linear program in equational form:*

$$\text{maximize } \mathbf{c}^T \mathbf{x} \text{ subject to } A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq 0$$

if there is at least one feasible solution, and the objective function is bounded above on $P(A, \mathbf{b})$, then there exists at least one optimal solution. Among the optimal solutions there is at least one basic solution.

PROOF. We claim that, for any feasible solution \mathbf{x}_0 , there exists a basic feasible solution \mathbf{x} with $\mathbf{c}^T \mathbf{x} \geq \mathbf{c}^T \mathbf{x}_0$. This implies that an optimal solution, if it exists, will be basic. Suppose \mathbf{x} is a feasible solution. Among all feasible solutions \mathbf{x} with $\mathbf{c}^T \mathbf{x} \geq \mathbf{c}^T \mathbf{x}_0$ choose one with support of minimal cardinality and call it $\tilde{\mathbf{x}}$. We will show that $A_{\text{supp}(\tilde{\mathbf{x}})}$ has linearly independent columns, so that $\tilde{\mathbf{x}}$ is basic.

Suppose this is not the case. Then there exists a vector $\mathbf{y} \in \mathbf{R}^n$ with $\text{supp}(\mathbf{y}) \subset \text{supp}(\tilde{\mathbf{x}})$ such that $A\mathbf{y} = 0$. Replacing \mathbf{y} by $-\mathbf{y}$ if necessary, assume that $\mathbf{c}^T \mathbf{y} \geq 0$.

We claim that we may further assume that \mathbf{y} has at least one *negative* coordinate. Suppose that all the coordinates of \mathbf{y} are non-negative. If $\mathbf{c}^T \mathbf{y} = 0$, then we can replace \mathbf{y} with $-\mathbf{y}$. If $\mathbf{c}^T \mathbf{y} > 0$ and all coordinates of \mathbf{y} are positive, then $\tilde{\mathbf{x}} + t\mathbf{y}$ is a feasible solution for all $t > 0$. The objective function $\mathbf{c}^T(\tilde{\mathbf{x}} + t\mathbf{y})$ grows unboundedly as t grows, contradicting its boundedness.

Thus \mathbf{y} has at least one negative coordinate, hence it is possible to choose a value $t > 0$ such that $\tilde{\mathbf{x}} + t\mathbf{y}$ is a feasible solution with $\text{supp}(\tilde{\mathbf{x}} + t\mathbf{y}) \subsetneq \text{supp}(\tilde{\mathbf{x}})$. Since $\mathbf{c}^T \mathbf{y} \geq 0$, we have $\mathbf{c}^T(\tilde{\mathbf{x}} + t\mathbf{y}) \geq \mathbf{c}^T \tilde{\mathbf{x}}$ and $\text{supp}(\tilde{\mathbf{x}} + t\mathbf{y}) \subsetneq \text{supp}(\tilde{\mathbf{x}})$. This contradicts the minimality condition on the cardinality of $\text{supp}(\tilde{\mathbf{x}})$.

The set of basic feasible solutions is finite. The element of this set that maximizes the objective function must therefore be an optimal solution. \square

DEFINITION 1.9 (Vertex). Let $P \subset \mathbf{R}^n$ be convex closed set. An element $\mathbf{v} \in P$ is said to be a *vertex* of P if there exists $\mathbf{c} \in \mathbf{R}^n$ such that $\mathbf{c}^T \mathbf{x}$ attains its maximum *uniquely* at \mathbf{v} .

Theorem 1.8 says that every vertex of $P(A, \mathbf{b})$ is a basic feasible solution. The converse is also true:

THEOREM 1.10. *The basic feasible solutions to (LP) are precisely the vertices of $P(A, \mathbf{b})$.*

PROOF. Let $B \subset [n]$ be a basic subset \mathbf{v} be the basic feasible solution to (LP) with respect to B . Define \mathbf{c} to be the vector with $c_j = 0$ for $j \in B$, and $c_j = -1$ (or any strictly negative number) otherwise. Then $\mathbf{c}^T \mathbf{v} = 0$, and by Lemma 1.6, $\mathbf{c}^T \mathbf{x} < 0$ for every $\mathbf{x} \in P(A, \mathbf{b})$ different from \mathbf{x}_0 . \square

DEFINITION 1.11 (General form of a linear program). A more general form of a linear program involves linear inequalities and equalities. As before take A to be an $m \times n$ matrix with real entries, $\mathbf{b} \in \mathbf{R}^m$, and $\mathbf{c} \in \mathbf{R}^n$. A general linear program has the form:

$$(\text{GLP}) \text{ optimize } \mathbf{c}^T \mathbf{x} \text{ subject to } a_{i1}x_1 + \cdots + a_{in}x_n \begin{matrix} R_i \\ S_j \end{matrix} b_i \text{ for } i = 1, \dots, m, x_j \begin{matrix} R_i \\ S_j \end{matrix} 0$$

where each R_i and S_j is one of the three symbols \leq , \geq , and $=$, and the word optimize is replaced by either maximize, or minimize. A basic feasible solution is one that is defined by equalities in n linearly independent constraints (which could be equality or inequality to begin with).

EXAMPLE 1.12 (Standard equational form of the simplex). Consider the linear program in n variables with just one linear equation:

$$\mathbf{x} \geq 0; x_1 + \cdots + x_n = 1.$$

The matrix A in this case has a single row, and rank one. The polytope $P(A, 1)$ is called the standard $(n - 1)$ -simplex. Every singleton subset of $[n]$ is basic. The basic solution corresponding to $B = \{j\}$ is the i th coordinate vector e_j . Given an objective vector $\mathbf{c} \in \mathbf{R}^n$, the optimal solution is e_j where j is any of the indices for which c_j is maximal among the coordinates of \mathbf{c} .

EXAMPLE 1.13. The cube can be defined by the inequalities:

$$0 \leq x_i \leq 1, \text{ for } i = 1, 2, 3.$$

The inequality $x_i \leq 1$ can be turned into an equality by introducing *slack variables* y_i , and writing:

$$x_i \geq 0, y_i \geq 0, x_i + y_i \leq 1 \text{ for } i = 1, 2, 3.$$

The linear program in equational form is equivalent to the original, more general one, in the sense that there is a bijection amongst their feasible solutions that maps vertices to vertices (why?). What are the basic subsets? What are the basic feasible solutions?

EXERCISE 1.14 (The simplex in terms of inequalities). The n -simplex can also be expressed in terms of inequalities:

$$\Delta_n = \{(x_1, \dots, x_n) \mid 0 \leq x_1 \leq x_2 \leq \cdots \leq x_n \leq 1\}.$$

Rewrite this in equational form. Determine the basic sets and basic solutions.

EXERCISE 1.15. Express the hyperoctahedron:

$$H_n = \{\mathbf{x} \mid -1 \leq x_1 + \cdots + x_n \leq 1\}$$

in equational form.

2. Simplex Tableaus

The simplex method begins with a basic set B for which there exists a basic feasible solution. Let $\bar{B} = [n] - B$, the complement of B in $[n]$. Using the relations imposed by $A\mathbf{x} = \mathbf{b}$, each of the basic variables x_j , $j \in B$, can be expressed in terms of the non-basic variables x_j , $j \in \bar{B}$. Using this, the objective function can also be expressed in terms of the non-basic variables.

To do this explicitly, note that the system of equations $A\mathbf{x} = \mathbf{b}$ can be rearranged as:

$$(2) \quad A_B \mathbf{x}_B = \mathbf{b} - A_{\bar{B}} \mathbf{x}_{\bar{B}}.$$

Here $\mathbf{x}_B \in \mathbf{R}^m$ and $\mathbf{x}_{\bar{B}} \in \mathbf{R}^{m-n}$ are vectors whose coordinates are those coordinates of \mathbf{x} whose indices lie in B and \bar{B} respectively. Since A_B is invertible, the basic variables can be expressed in terms of the non-basic ones:

$$\mathbf{x}_B = A_B^{-1}(\mathbf{b} - A_{\bar{B}} \mathbf{x}_{\bar{B}}).$$

Indeed, the basic feasible solution is computed by setting $\mathbf{x}_{\bar{B}} = 0$ in the above equation. If $A_B^{-1}\mathbf{b} \geq 0$, then it is the basic feasible solution corresponding to B . Otherwise there is no basic feasible solution corresponding to B . Since each basic

variable is expressed in terms of the non-basic variables in (2), the objective function can be expressed in terms of the non-basic variables only:

$$\mathbf{c}^T \mathbf{x} = \mathbf{c}_B^T \mathbf{x}_B + \mathbf{c}_{\bar{B}}^T \mathbf{x}_{\bar{B}} = \mathbf{c}_B^T A_B^{-1} (\mathbf{b} - A_{\bar{B}} \mathbf{x}_{\bar{B}}) + \mathbf{c}_{\bar{B}}^T \mathbf{x}_{\bar{B}}.$$

Thus (LP) is represented in terms of a *simplex tableau*:

$$(T) \quad \begin{array}{rcl} \mathbf{x}_B & = & \mathbf{d} - D \mathbf{x}_{\bar{B}} \\ \hline \mathbf{c}^T \mathbf{x} & = & e - \mathbf{e}^T \mathbf{x}_{\bar{B}}, \end{array}$$

the equations above the line being the constraints, and the objective function $\mathbf{c}^T \mathbf{x}$ below the line to be maximized. Here:

$$\begin{aligned} \mathbf{d} &= A_B^{-1} \mathbf{b} \\ D &= A_B^{-1} A_{\bar{B}} \\ e &= \mathbf{c}_B^T \mathbf{d} \\ \mathbf{e}^T &= \mathbf{c}_B^T D - \mathbf{c}_{\bar{B}}^T. \end{aligned}$$

The information contained in (T) is equivalent to the information in (LP). But (T) gives a parametrization of $P(A, \mathbf{b})$ in terms of a subset of \mathbf{R}^{n-m} .

THEOREM 2.1. *For every basic subset $B \subset [n]$ in the linear program (LP),*

$$(*) \quad P(A, \mathbf{b}) = \{\mathbf{x} \in \mathbf{R}^n \mid \mathbf{x}_B = A_B^{-1} (\mathbf{b} - A_{\bar{B}} \mathbf{x}_{\bar{B}}) \geq 0, \mathbf{x}_{\bar{B}} \geq 0\}.$$

In other words, (T) gives a parametrization of $P(A, \mathbf{b})$ in terms of the polytope:

$$P_B(A, \mathbf{b}) = \{\mathbf{x}_{\bar{B}} \in \mathbf{R}_{\geq 0}^{n-m} \mid A_B^{-1} (\mathbf{b} - A_{\bar{B}} \mathbf{x}_{\bar{B}}) \geq 0\} \subset \mathbf{R}^{n-m}.$$

PROOF. The conditions $A\mathbf{x} = \mathbf{b}$ and $\mathbf{x}_B = A_B^{-1} (\mathbf{b} - A_{\bar{B}} \mathbf{x}_{\bar{B}})$ are equivalent. \square

EXAMPLE 2.2. Consider the Birkhoff polytope (Example 1.3) for $r = 3$. For convenience we abbreviate the variable indices (i, j) to ij . The matrix A has columns indexed by pairs in $[d] \times [d]$ written in increasing lexicographic order:

$$A = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \end{pmatrix}$$

Fix as objective function $x_{11} + x_{22} + x_{33}$. A basic subset is $B = \{11, 12, 13, 22, 31\}$ with basic solution given by $x_{13} = x_{22} = x_{31} = 1$, and all other coordinates zero. The corresponding simplex tableau is:

$$\begin{array}{rcl} x_{11} & = & -x_{21} + x_{32} + x_{33} \\ x_{12} & = & x_{21} + x_{23} - x_{32} \\ x_{13} & = & 1 - x_{23} - x_{33} \\ x_{22} & = & 1 - x_{21} - x_{23} \\ x_{31} & = & 1 - x_{32} - x_{33} \\ \hline \mathbf{c}^T \mathbf{x} & = & 1 - 2x_{21} - x_{23} + x_{32} + 2x_{33}. \end{array}$$

In the above example, the objective function can be increased by increasing x_{32} or x_{33} , the variables with positive coefficients in the last row of the tableau. However, this increase should respect the constraints that all the variables are non-negative. The condition $x_{12} \geq 0$ (using the second equation, and leaving the values of x_{21} and x_{23} unchanged at 0) gives $x_{32} \leq 0$. Therefore, it is not feasible to increase

x_{32} . However, it is feasible to increase x_{33} . The conditions $x_{13} \geq 0$ and $x_{31} \geq 0$ give $x_{33} \leq 1$. So we set $x_{33} = 1$, and recalculate all the basic variables, getting $x_{11} = x_{22} = x_{33} = 1$, and all other variables 0. We move x_{33} to the set of basic variables, and move x_{13} to the set of non-basic (which has now become 0), and use the equation:

$$x_{33} = 1 - x_{13} - x_{23}.$$

Using this we get a new tableau:

$$\begin{array}{rcl} x_{11} & = & 1 - x_{21} + x_{32} + x_{23} - x_{13} \\ x_{12} & = & x_{21} + x_{23} - x_{32} \\ x_{22} & = & 1 - x_{21} - x_{23} \\ x_{31} & = & x_{13} + x_{23} - x_{32} - x_{33} \\ x_{33} & = & 1 - x_{23} - x_{13} \\ \hline \mathbf{c}^T \mathbf{x} & = & 3 - 2x_{13} - 2x_{21} - 3x_{23} + x_{32}. \end{array}$$

All the non-basic variables have negative coefficients, except x_{32} . However, the constraint $x_{12} \geq 0$ still does not allow us to increase x_{32} without changing any other non-basic variable. This suggests that we may have arrived at a maximum value for the objective function. Indeed, $x_{12} = x_{21} + x_{23} - x_{32} \geq 0$ implies that $x_{32} \leq x_{21} + x_{23}$ for every $\mathbf{x} \in P(A, \mathbf{b})$, whence

$$\mathbf{c}^T \mathbf{x} = 3 - 2x_{13} - 2x_{21} - 3x_{23} + x_{32} \leq 3 - 2x_{13} - x_{21} - 2x_{23} \leq 3.$$

Therefore 3 is indeed a global maximum for the objective function, and is obtained uniquely at $x_{ij} = \delta_{ij}$.

An alternative approach would be to induct x_{32} into the set of basic variables, and remove x_{12} . Now the basic set is changed to $\{11, 22, 31, 32, 33\}$, but the basic feasible solution remains unchanged. This will result in the tableau:

$$\begin{array}{rcl} x_{11} & = & 1 - x_{12} + x_{32} + 2x_{23} - x_{13} \\ x_{22} & = & 1 - x_{21} - x_{23} \\ x_{31} & = & x_{12} + x_{13} - x_{21} - x_{33} \\ x_{32} & = & x_{21} + x_{23} - x_{12} \\ x_{33} & = & 1 - x_{23} - x_{13} \\ \hline \mathbf{c}^T \mathbf{x} & = & 3 - 2x_{13} - x_{21} - 2x_{23} - x_{12}. \end{array}$$

In this case, all the coefficients of the objective function, when expressed in terms of non-basic variables, are negative. Theorem 2.1 then implies that $x_{ij} = \delta_{ij}$ is the unique global maximum.

More generally Theorem 2.1 gives:

THEOREM 2.3. *Let B be a basic set for (LP). Suppose that the objective function, when expressed in terms of the non-basic variables as in the last line of the tableau (T) has all coefficients nonpositive. Then the basic feasible solution \mathbf{x}_0 corresponding to B is a solution to (LP). If all the non-basic variables occur with strictly negative coefficients, then \mathbf{x}_0 is the unique solution to (LP).*

3. Pivot Rules

Given a basic subset B for the linear program (LP), consider the corresponding tableau (T). Assume also that B admits a basic feasible solution \mathbf{x} . Suppose that

$e_k > 0$ for some $k \in \bar{B}$. For each i in B , we have:

$$(3) \quad x_i = d_i - \sum_{j \in \bar{B}} D_{ij} x_j.$$

The constraint $x_i \geq 0$ can be rewritten as:

$$D_{ik} x_k \leq d_i - \sum_{j \in \bar{B}, j \neq k} D_{ij} x_j.$$

Since B admits a basic feasible solution where all the basic variables are zero, we have $d_i \geq 0$. If $D_{ik} \leq 0$, then x_k can be increased without bound without violating (3). If, on the other hand, $D_{ik} > 0$, then x_k can be increased up to d_i/D_{ik} . Therefore, we can increase x_k up to $m = \min\{d_i/D_{ik} \mid i \in B, D_{ik} > 0\}$. In general $m \geq 0$, and it is quite possible (as we have seen in Example 2.2) that $m = 0$. Let l be any index for which $m = d_l/D_{lk}$. Let $B' = B \cup \{k\} - \{l\}$.

We have an expression for x_k in terms of the variables $\{x_j \mid j \in \bar{B}'\}$:

$$(4) \quad x_k = D_{lk}^{-1}(d_l - x_l - \sum_{j \in \bar{B}, j \neq k} D_{lj} x_j).$$

This can be substituted into the equations (3) for all $i \in B - \{l\}$ to write x_i in terms of the variables $\{x_j \mid j \in \bar{B}'\}$. Thus each variable x_i , for $i \in B'$ can be expressed in terms of the variables $\{x_j \mid j \in \bar{B}'\}$. Also, the substitution (4) can be used to express the objective function in terms of the variables $\{x_j \mid j \in \bar{B}'\}$. Let \mathbf{x}' be the vector obtained from \mathbf{x} by changing x_k to m , and x_l to 0.

LEMMA 3.1. *If B is a basic subset with basic feasible vector \mathbf{x} and B' and \mathbf{x}' are obtained from B and \mathbf{x} as described above, then B' is also a basic subset, and \mathbf{x}' is a basic feasible solution corresponding to B' .*

PROOF. Since B is a basic subset, the analysis at the beginning of Section 2 shows that the equations:

$$A\mathbf{x} = \mathbf{b} \text{ and } \mathbf{x}_B = \mathbf{d} - D\bar{\mathbf{x}}_{\bar{B}}$$

are equivalent. Our choices of k and l above ensure that these equations are further equivalent to

$$\mathbf{x}_{B'} = \mathbf{d}' - D'\bar{\mathbf{x}}_{\bar{B}'},$$

for some vector \mathbf{d}' and some matrix D' . Setting $\mathbf{x}_{\bar{B}'} = 0$ says that there is a unique value for $\mathbf{x}_{B'}$ such that, if $\mathbf{x}'' \in \mathbf{R}^n$ is obtained from $\mathbf{x}_{B'}$ by setting the coordinates in \bar{B}' to 0, then $A\mathbf{x}'' = \mathbf{b}$. In other words, $A_{B'}\mathbf{x}_{B'} = \mathbf{b}$ has a unique solution. It follows that $A_{B'}$ is non-singular, and so B' basic. By Lemma 1.6, the vector \mathbf{x}'' constructed in this proof coincides with the vector \mathbf{x}' constructed earlier, so the rest of the lemma also follows. \square

DEFINITION 3.2 (Pivot step). Given a basic subset $B \subset [n]$ and the corresponding tableau (T) for a linear program (LP), a *leaving variable* is any variable x_k such that $k \in \bar{B}$, and $e_k > 0$. Let $m = \min\{d_i/D_{ik} \mid D_{ik} > 0\}$. Assume that $m < \infty$ (we will come to the case $m = \infty$ shortly). Having picked a leaving variable x_k , an *entering variable* is any variable x_l such that $m = d_l/D_{lk}$. A pivot step is a change of basic set $B \mapsto B' := B \cup \{x_l\} - \{x_k\}$, where x_l is an entering variable and x_k is a leaving variable. By Theorem 2.3, B' is again a basic subset for (LP). If \mathbf{x} is the basic feasible solution corresponding to B , and \mathbf{x}' is the basic feasible solution corresponding to B' , then $\mathbf{c}^T \mathbf{x}' \geq \mathbf{c}^T \mathbf{x}$.

In what situation is there no possible pivot step? The first possibility is that $e_k \leq 0$ for all $k \in \bar{B}$. Then Theorem 2.3 tells us that the basic feasible solution corresponding to B maximizes the objective function of $P(A, \mathbf{b})$. Another possibility is that, for some $k \in \bar{B}$ such that $e_k > 0$, $D_{ik} < 0$ for all $i \in B$. Then, starting with the basic feasible solution corresponding to B , x_k can be increased unboundedly without disturbing any of the other non-basic variables (which are set to 0) without making any of the variable x_i , $i \in B$ negative. This shows that the objective function $\mathbf{c}^T \mathbf{x}$ is unbounded on $P(A, \mathbf{b})$. If neither of these cases occur, then a pivot step $B \mapsto B'$ is possible. However, it is quite possible that $m = 0$ (this happened with the variable x_{32} in Example 2.2). In this case the pivot step $B \mapsto B'$ happens, but $\mathbf{x}' = \mathbf{x}$. Of course, the value of the objective function remains unchanged. Also, $\text{supp}(\mathbf{x}) = \text{supp}(\mathbf{x}') \subset B \cap B' \subsetneq B$.

DEFINITION 3.3 (Degenerate basic subset). A basic subset for (LP) is said to be *degenerate* if it admits a basic feasible solution \mathbf{x} such that $\text{supp}(\mathbf{x}) \subsetneq B$.

LEMMA 3.4. *Given an $m \times n$ matrix A , the set of constraint vectors $\mathbf{b} \in \mathbf{R}^n$ for which there exist degenerate basic subsets is a finite union of hyperplanes.*

PROOF. If $B \subset [n]$ is degenerate, then the vector $A_B^{-1} \mathbf{b}$ lies in the union of coordinate hyperplanes. In other words, \mathbf{b} lies in image under A_B of a coordinate hyperplane, which, by the non-singularity of A_B , is a hyperplane. Thus the set of constraint vectors \mathbf{b} for which there exists a degenerate basic subset is a finite union of a hyperplanes. \square

EXAMPLE 3.5 (Degeneracy in transportation polytopes). When are the defining equations for $\mathbf{M}_{\mathbf{ab}}$ from Example 1.2 degenerate? Suppose that \mathbf{x} is a basic feasible solution corresponding to a basic subset B . If $\text{supp}(\mathbf{x}) \subsetneq B$, then $\Gamma_{\text{supp} \mathbf{x}}$ is a forest with at least two connected components. Therefore there exist proper subsets $K \subset [r]$ and $L \subset [s]$ with complements $[r] - K$ and $[s] - L$ respectively such that $\Gamma_{\text{supp} \mathbf{x}}$ has no edges between K and \bar{L} or between \bar{K} and L . It follows that

$$\sum_{i \in K} a_i = \sum_{i \in K} \sum_{j \in L} x_{ij} = \sum_{j \in L} b_j.$$

Conversely suppose there exist proper subset $K \subset [r]$ and $L \subset [s]$ such that

$$(5) \quad \sum_{i \in K} a_i = \sum_{j \in L} b_j,$$

then (as we shall see in Theorem 3.11) both $\mathbf{M}_{\mathbf{a}_K \mathbf{b}_L}$ and $\mathbf{M}_{\mathbf{a}_{\bar{K}} \mathbf{b}_{\bar{L}}}$ are non-empty. Putting together feasible points for these polytopes will give us a point \mathbf{x} in $\mathbf{M}_{\mathbf{ab}}$ whose support is a forest, hence a proper subset of any basic subset. Thus we have the following:

THEOREM 3.6. *Assume that $\mathbf{a} > 0$ and $\mathbf{b} > 0$. The polytope $\mathbf{M}_{\mathbf{ab}}$ is degenerate if and only if there exist proper subsets $K \subset [r]$ and $L \subset [s]$ such that the condition (5) holds.*

Since a finite union of coordinate hyperplanes is a nowhere dense subset of \mathbf{R}^n , it is very unlikely that degeneracy will occur in a real-world linear programming problem. Moreover, if it does occur, a small enough perturbation of \mathbf{b} will bring us back to the non-degenerate world while giving the correct basic subset whose basic feasible solution also optimizes the original problem.

DEFINITION 3.7 (Non-degenerate linear program). We say that (LP) is non-degenerate if it admits no degenerate basic subset $B \subset [n]$.

THEOREM 3.8. Assume that (LP) is a non-degenerate linear program. Starting with any basic set B , a finite number of pivot steps (Definition 3.2) will result in a basic set whose basic feasible solution is optimal, or else establish that the objective function is unbounded on $P(A, \mathbf{b})$.

PROOF. Indeed, non-degeneracy ensures that the objective function increases strictly at each pivot step or $m = \infty$. If $m = \infty$ at any pivot step, then the objective function is unbounded. Otherwise, because the objective function increases strictly, no basic subset is ever repeated in a run of the algorithm. Since there are only finitely many possible basic subsets, the algorithm must halt at an optimal solution. \square

The number of basic subsets can grow exponentially with the number of variable. In an implementation of a linear programming algorithm a rule for choosing the leaving variable and the entering variable has to be specified. Such a rule is called a *pivot rule*. For a discussion of commonly used pivot rules, their relative merits and demerits, the reader is referred to [1, Section 5.7].

However, degeneracy is ubiquitous in combinatorial applications of linear programming. For example, *every* basic subset for the r th Birkhoff polytope for every $r > 1$ either does not admit a basic feasible solution, or is degenerate, since every vertex of the Birkhoff polytope is a permutation matrix. Moreover, the support of any permutation matrix has r elements, but every basic subset has cardinality $2r - 1$.

EXAMPLE 3.9. Consider the Birkhoff polytope for $r = 3$ with objective function:

$$x_{12} + x_{13} + x_{21} + x_{23} + x_{31} + x_{32}.$$

For $B = \{11, 12, 13, 22, 33\}$, the tableau is

$$\begin{array}{rcl} x_{11} & = & 1 - x_{21} - x_{31} \\ x_{12} & = & x_{21} + x_{23} - x_{32} \\ x_{13} & = & -x_{23} + x_{31} + x_{32} \\ x_{22} & = & 1 - x_{21} - x_{23} \\ x_{33} & = & 1 - x_{31} - x_{32} \\ \hline \mathbf{c}^T \mathbf{x} & = & 2x_{21} + x_{23} + 2x_{31} + x_{32}. \end{array}$$

All four non-basic variables have positive coefficients in the objective function. Using x_{23} or x_{32} as entering variables leads to pivots that do not increase the objective function.

The simplex method begins with a basic subset $B \subset [n]$ for (LP). How does one find a basic subset to begin with? Indeed, the total number of subsets of $[n]$ is 2^n , and finding a basic subset could be like finding a needle in a (albeit finite) haystack. The trick is to start with an infeasible point (usually $\mathbf{x} = 0$), quantify its “infeasibility” and then move to a feasible point (if it exists) by taking the negative of this infeasibility as an objective function.

So, introduce variables x_{n+1}, \dots, x_{n+m} , all constrained to take non-negative values. The variable x_{n+i} will measure the feasibility gap in the i th equation:

$$a_{i1}x_1 + \dots + a_{in}x_n + \text{sgn}(b_i)x_{n+i} = b_i,$$

and then maximize $-x_{n+1} - x_{n+1} - \dots - x_{n+m}$. In other words, the auxilliary linear program has coefficient matrix $\tilde{A} = (A \ I_m)$ in block form, where I_m denotes the $m \times m$ identity matrix. The bound vector \mathbf{b} remains unchanged. The objective vector is $(0, \dots, 0, -1, \dots, -1)$ (with n zeros and m -1 's). In this case $B = \{n+1, \dots, n+m\}$ is clearly a basic subset with basic feasible solution $\tilde{\mathbf{x}} = (0, \dots, 0, |b_1|, \dots, |b_m|)$. The original program (LP) has a feasible solution if and only if the maximum value of the objective function is 0. Thus locating a basic subset is itself a linear programming problem with an obvious basic subset.

EXAMPLE 3.10 (Construction of transportation arrays). In order to construct an $m \times n$ transportation array with margins \mathbf{a} and \mathbf{b} (Example 1.2), we introduce $r + s$ auxilliary variables u_1, \dots, u_r and v_1, \dots, v_s . Consider the auxilliary linear program:

$$(6) \quad \begin{array}{rcl} \sum_{j=1}^s x_{ij} + u_i & = & a_i \text{ for } 1 \leq i \leq r, \\ \sum_{i=1}^r x_{ij} + v_j & = & b_j \text{ for } 1 \leq j \leq s, \\ x_{ij} \geq 0, u_i \geq 0, v_j \geq 0 & = & \text{for } 1 \leq i \leq r, 1 \leq j \leq s, \\ \hline & \text{maximize} & -u_1 - \dots - u_r - v_1 - \dots - v_s. \end{array}$$

This auxilliary program comes with a God-given basic variables $\{u_1, \dots, u_r, v_1, \dots, v_s\}$. The corresponding tableau is:

$$\begin{array}{rcl} u_i & = & a_i - \sum_{j=1}^s x_{ij} \text{ for } i = 1, \dots, r, \\ v_j & = & b_j - \sum_{i=1}^r x_{is} \text{ for } j = 1, \dots, s, \\ \hline \mathbf{c}^T \mathbf{x} & = & -\sum a_i - \sum_j b_j + \sum_i \sum_j x_{ij}. \end{array}$$

We may choose any x_{ij} as entering variable. Then the leaving variable will be u_i if $a_i < b_j$, and v_j if $b_i < u_j$. If $a_i = b_j$ either variable could be chosen as leaving. We now have a new tableau, for example if we had chosen $x_{i_0 j_0}$ as entering u_{i_0} as leaving,

$$\begin{array}{rcl} x_{i_0 j_0} & = & a_{i_0} - u_{i_0} - \sum_{j \neq j_0} x_{i_0 j} \\ v_{j_0} & = & b_{j_0} - \sum_{i \neq i_0} x_{i j_0} - a_{i_0} + u_{i_0} + \sum_{j \neq j_0} x_{i_0, j} \\ u_i & = & a_i - \sum_j x_{ij} \text{ for } i \neq i_0, \\ v_j & = & b_j - \sum_{i=1}^r x_{is} \text{ for } j \neq j_0, \\ \hline \mathbf{c}^T \mathbf{x} & = & -\sum_{i \neq i_0} a_i - \sum b_j + \sum_{i \neq i_0} \sum_j x_{ij} - u_{i_0}. \end{array}$$

The variables $x_{i_0, j}$ (the i_0 th row) have dissappeared from the objective function. The i_0 th row is now saturated and the rest of the algorithm would run with the same set of possible pivot steps as if the i th row simply did not exist, and $b_{j_0} = b_j - a_{i_0}$. The new margin vectors are:

$$\mathbf{a}' = (a_1, \dots, a_{i_0-1}, a_{i_0+1}, \dots, a_r), \quad \mathbf{b}' = (b_1, \dots, b_{j_0-1}, b_{j_0} - a_{i_0}, b_{j_0+1}, \dots, b_s).$$

We now analyze the running of this auxilliary program by induction on $r + s$.

THEOREM 3.11. *The transportation polytope $\mathbf{M}_{\mathbf{ab}}$ is non-empty if and only if $\sum_{i=1}^r a_i = \sum_{j=1}^s b_j$. If this condition holds, then every basic feasible solution in $\mathbf{M}_{\mathbf{ab}}$ can be obtained by a sequence of pivot steps for the auxilliary linear program described above.*

The method for finding all the vertices of $\mathbf{M}_{\mathbf{ab}}$ in the above theorem is equivalent to the algorithm of Corollary 8.1.4 in Brualdi's book [3].

PROOF. Clearly if $\mathbf{M}_{\mathbf{ab}}$ has a feasible point $\mathbf{x} = (x_{ij})$ then $\sum a_i = \sum_i \sum_j = \sum_j b_j$. For the converse, proceed by induction on $r + s$. Take as base case $r + s = 2$, where the array has one row and one column, and clearly it is feasible if the row margin equals the column margin. As we have seen, each pivot step reduces $r + s$ by one, and reduces both the margins by the same quantity. Thus by the induction hypothesis the new transportation polytope is non-empty.

It remains to show that every basic feasible solution can be attained by a sequence of pivot steps. Suppose that \mathbf{x} is a basic feasible solution. By Example 1.4, $\Gamma_{\text{supp}\mathbf{x}}$ is contained in a spanning tree, and hence is a spanning forest. Consider a leaf of this graph, say a row index $1 \leq i_0 \leq r$ (the case where it is a column index is not different). Being a leaf, this vertex is connected to a unique column index $1 \leq j_0 \leq s$. Perform the first pivot step for the auxiliary program taking $x_{i_0 j_0}$ as entering variable and u_{i_0} as leaving variable. As discussed before what remains is a problem with $r - 1$ rows and s columns, for which all basic feasible solutions can indeed be obtained by induction hypothesis. \square

THEOREM 3.12 (Birkhoff von Neumann theorem). *The vertices of the Birkhoff polytope are permutation matrices.*

PROOF. Every permutation matrix, being supported on a forest (in fact a perfect matching) of $K_{n,n}$ is a vertex. The converse follows from observing (by induction) that every run of the algorithm of Example 3.10 produces a permutation matrix. \square

4. Dealing with degeneracy

In the non-degenerate case, we saw that, with an sequence of arbitrarily chosen pivot moves, it is possible to go from any basic feasible solution to an optimal one. In the degenerate case, it is quite possible that we end up repeating a cycle of pivot moves indefinitely. This can be avoided using *Bland's rule* for choosing the leaving and entering variable.

DEFINITION 4.1 (Bland's rule). Given a basic subset $B \subset [n]$ for (LP), choose the incoming variable with minimum possible index, and among all the outgoing variables determined by this incoming variable, choose the one with minimum possible index.

THEOREM 4.2. *With Bland's rule, starting with a given basic subset $B \subset [n]$ for (LP), it is not possible to return to B after a any number of pivot steps.*

PROOF. Suppose that, upon using pivot steps according to Bland's rule, the simplex method encounters a cycle. Let F denote the set of all the entering (or leaving variables) in the pivot steps of this cycle. Since the objective function does not increase in a cycle, it must be the case that the basic feasible solutions do not change. Therefore for all basic feasible solutions in the cycle $x_j = 0$ for all $j \in F$.

Let v be the largest index in F . Let B be a basic subset in the cycle just before x_v enters the basic set. Suppose that the simplex tableau corresponding to B is:

$$\begin{aligned} \frac{\mathbf{x}_B}{\mathbf{c}^T \mathbf{x}} &= \frac{\mathbf{d} - D\mathbf{x}_{\bar{B}}}{\mathbf{e} - \mathbf{e}^T \mathbf{x}_{\bar{B}}}, \end{aligned}$$

Since we are following Bland's rule, v is the least index among all possible entering variables. Therefore for any $j \in F - \{v\}$, since $j < v$, $e_j \leq 0$.

Let B' be a basic subset in the cycle just before x_v leaves the basic set. Suppose that the simplex tableau corresponding to B' is:

$$\frac{\mathbf{x}_{B'} = \mathbf{d}' - D' \mathbf{x}_{\bar{B}'}}{\mathbf{c}^T \mathbf{x} = e' - \mathbf{e}'^T \mathbf{x}_{\bar{B}'},}$$

Suppose x_l enters the basic set as v leaves it. Now for all $j \in F$ d_j , being the value of x_j in the basic feasible solution corresponding to B' , must be 0. Since we are following Bland's rule, v is the least index among all possible leaving variables. Therefore, for any $j \in F - \{v\}$, since $j < v$ and $d'_v/D'_{vk'} = d'_j/D'_{jk'} = 0$, we must have $D'_{jk'} < 0$.

To summarize, we have established:

$$(7) \quad e_v > 0 \text{ and for all } j \in F - \{v\}, e_j \leq 0,$$

$$(8) \quad D_{vk'} > 0 \text{ and for all } j \in F - \{v\}, D'_{jk'} < 0.$$

Now consider the auxilliary linear program:

$$\text{maximize } \mathbf{c}^T \mathbf{x} \text{ subject to } \mathbf{A}\mathbf{x} = \mathbf{b}, x_v \leq 0, \mathbf{x}_{F-\{v\}} \geq 0, \mathbf{x}_{\bar{B}-F} = 0.$$

There is no constraint on the variables \mathbf{x}_{B-F} . They may assume positive or negative signs, or take the value 0.

Let $\tilde{\mathbf{x}}$ denote the basic feasible solution to (LP) corresponding to B . It is also feasible for the auxilliary program. The condition (7) and the constraints on the auxilliary program imply that, for all feasible \mathbf{x} in the auxilliary program, $e_j \mathbf{x}_j \leq 0$ for all $j \in \bar{B}$. Therefore, $\tilde{\mathbf{x}}$ is optimal.

Now $\tilde{\mathbf{x}}$ is also the basic feasible solution to the auxilliary program corresponding to the basic subset B' . Since there are no sign constraints on \mathbf{x}_{B-F} , the conditions (8) imply that, starting with $\tilde{\mathbf{x}}$ the leaving variable $x_{k'}$ can be increased indefinitely without violating feasibility. Therefore the objective function is unbounded in the auxilliary program. This contradicts the earlier assertion that the auxilliary program had an optimal solution. \square

5. Duality

Consider a linear program in the form:

$$(P) \quad \text{maximize } \mathbf{c}^T \mathbf{x} \text{ subject to } \mathbf{A}\mathbf{x} \leq \mathbf{b}, \mathbf{x} \geq 0,$$

with n variables $\mathbf{x} = (x_1, \dots, x_n)$, $A \in M_{m \times n}(\mathbf{R})$, $\mathbf{b} \in \mathbf{R}^m$. We will call (P) the *primal* program.

The program *dual* to (P) is defined as:

$$(D) \quad \text{minimize } \mathbf{b}^T \mathbf{y} \text{ subject to } A^T \mathbf{y} \geq \mathbf{c}, \mathbf{y} \geq 0,$$

with n variables $\mathbf{y} = (y_1, \dots, y_n)$.

Suppose \mathbf{x} and \mathbf{y} are feasible solutions to (P) and (D) respectively. Then, combining the constraints $\mathbf{c}^T \leq \mathbf{y}^T A$ and $\mathbf{A}\mathbf{x} \leq \mathbf{b}$, along with $\mathbf{x} \geq 0$ and $\mathbf{y} \geq 0$ gives:

$$\mathbf{c}^T \mathbf{x} \leq \mathbf{y}^T A \mathbf{x} \leq \mathbf{y}^T \mathbf{b}.$$

Thus if (P) and (D) are feasible, then the maximum of (P) is bounded above by the minimum of (D). It follows that if the objective function of (P) is unbounded above on its feasible set, then (D) has no feasible solutions. Similarly, if the objective function of (D) is unbounded below on its feasible set, then (P) has no feasible solutions.

THEOREM 5.1. *If (P) and (D) are both feasible, then the maximum value of (P) equals the minimum value of (D).*

PROOF. Convert (P) to equational form by introducing slack variables x_{m+1}, \dots, x_{m+n} . Let $\bar{\mathbf{x}} = (x_1, \dots, x_{m+n})$, $\bar{A} = (A \quad I_m)$, and $\bar{\mathbf{c}} = (c_1, \dots, c_n, 0, \dots, 0) \in \mathbf{R}^{m+n}$. The equational form of (P) is

$$\text{maximize } \bar{\mathbf{c}}^T \bar{\mathbf{x}} \text{ subject to } \bar{A} \bar{\mathbf{x}} = \mathbf{b}, \bar{\mathbf{x}} \geq 0.$$

Let $B \subset [m+n]$ be a basic subset whose basic feasible solution $\bar{\mathbf{x}}^*$ maximizes the objective function. The corresponding simplex tableau is:

$$(9) \quad \frac{\bar{\mathbf{x}}_B}{\bar{\mathbf{c}}^T \bar{\mathbf{x}}} = \frac{\bar{A}_B^{-1} \mathbf{b} - \bar{A}_B^{-1} \bar{A}_{\bar{B}} \bar{\mathbf{x}}_{\bar{B}}}{\bar{\mathbf{c}}^T \bar{A}_B^{-1} \mathbf{b} + (\bar{\mathbf{c}}_{\bar{B}}^T - \bar{\mathbf{c}}_B^T \bar{A}_B^{-1} \bar{A}_{\bar{B}}) \bar{\mathbf{x}}_{\bar{B}}}.$$

The optimal solution $\bar{\mathbf{x}}^*$ is obtained by setting $\bar{\mathbf{x}}_{\bar{B}} = 0$ in the above tableau. So the maximum value of the objective function is:

$$\bar{\mathbf{c}}^T \bar{\mathbf{x}}^* = \bar{\mathbf{c}}^T \bar{A}_B^{-1} \mathbf{b}.$$

So if $\mathbf{y}^* = (\bar{\mathbf{c}}_B^T \bar{A}_B^{-1})^T$, then

$$\mathbf{b}^T \mathbf{y}^* = \bar{\mathbf{c}}^T \bar{\mathbf{x}}^*.$$

We claim that \mathbf{y}^* is feasible for (D), i.e., $\mathbf{y}^* \geq 0$ and $A^T \mathbf{y}^* \geq \mathbf{c}$. These two inequalities can be combined and written as $\bar{A}^T \mathbf{y}^* \geq \bar{\mathbf{c}}$. This combined inequality can now be split up along rows from B and rows from \bar{B} :

$$\bar{A}_B^T \mathbf{y}^* \geq \bar{\mathbf{c}}_B \text{ and } \bar{A}_{\bar{B}}^T \mathbf{y}^* \geq \bar{\mathbf{c}}_{\bar{B}}.$$

Plugging in the value of \mathbf{y}^* into these inequalities turns them into:

$$\bar{A}_B^T (\bar{\mathbf{c}}_B^T \bar{A}_B^{-1})^T \geq \bar{\mathbf{c}}_B \text{ and } \bar{A}_{\bar{B}}^T (\bar{\mathbf{c}}_B^T \bar{A}_B^{-1})^T \geq \bar{\mathbf{c}}_{\bar{B}}.$$

The inequality on the left holds because it is actually an equality! As for the inequality on the right, it is just the non-positivity of the coefficients of the objective function when it is expressed in terms of $\bar{\mathbf{x}}_{\bar{B}}$ as in the tableau (9) (Theorem 2.3).

To prove the converse, observe that (D) can be written as:

$$\text{maximize } -\mathbf{b}^T \mathbf{y} \text{ subject to } -A^T \mathbf{y} \leq -\mathbf{c} \text{ and } \mathbf{y} \geq 0.$$

The dual of this program is:

$$\text{minimize } -\mathbf{c}^T \mathbf{x} \text{ subject to } -A\mathbf{x} \geq -\mathbf{b} \text{ and } \mathbf{x} \geq 0,$$

which is equivalent to (P). \square

Duality applies equally well to more general linear programs of the form (GLP) in Example 1.11. For example, a program in equational form (LP) can be rewritten as:

$$\text{maximize } \mathbf{c}^T \mathbf{x} \text{ subject to } A\mathbf{x} \leq \mathbf{b}, -A\mathbf{x} \leq -\mathbf{b}, \mathbf{x} \geq 0.$$

The resulting dual program will have $2m$ variables (where m is the number of rows of A), say $\mathbf{u} = u_1, \dots, u_{2m}$. The dual program is:

$$\text{minimize } (\mathbf{b}^T \quad -\mathbf{b}^T) \mathbf{u} \text{ subject to } (A^T \quad A^T) \mathbf{u} \geq \mathbf{c} \text{ and } \mathbf{u} \geq 0.$$

Since the i th column of the constraint matrix $\tilde{A} = (A^T \quad A^T)$ coincides with its $n+i$ th column a basic subset can have one or the other, but not both. Let $y_i = u_i - u_{n+i}$. Then $\mathbf{b}^T \mathbf{y} = (\mathbf{b}^T \quad -\mathbf{b}^T) \mathbf{u}$. However, the variables y_i can be negative as well as non-negative. Thus the dual program is:

$$\text{minimize } \mathbf{b}^T \mathbf{y} \text{ subject to } A^T \mathbf{y} \geq \mathbf{c}.$$

Unlike (D), the constraint $\mathbf{y} \geq 0$ is missing.

Thinking of the relations occurring in (GLP) as a vectors $R = (R_1, \dots, R_m)$ and $S = (S_1, \dots, S_n)$ of relations, each R_i taken from the set $\{\geq, \leq, =\}$, and each S_j taken from $\{\geq, \leq, \geq \text{ or } \leq\}$, we write (GLP) as

$$\text{(GLP)} \quad \text{Ximize } \mathbf{c}^T \mathbf{x} \text{ subject to } A\mathbf{x} R \mathbf{b}, \mathbf{x} S \mathbf{0}.$$

the dual is given by

$$\text{(GDP)} \quad \text{Yimize } \mathbf{b}^T \mathbf{y} \text{ subject to } A^T \mathbf{y} S' \mathbf{c} \text{ and } \mathbf{y} R' \mathbf{0}$$

where if X is max then Y is min and vice-versa, if R_i is $\geq, \leq, =$, then R'_i is $\leq, \geq, \geq \text{ or } \leq$, respectively, and if S_j is $\geq, \leq, \geq \text{ or } \leq$, then S'_j is $\geq, \leq, =$, respectively.

THEOREM 5.2 (Duality theorem for general linear programs). *If (GLP) is bounded and feasible, then so is (GDP), and their optimal values are equal.*

EXERCISE 5.3. Construct an example of an infeasible linear program of the form (P) such that the dual program (D) is also infeasible.

COROLLARY 5.4 (Farkas lemma). *Let A be an $m \times n$ real matrix and $\mathbf{b} \in \mathbf{R}^m$ be a vector. Then exactly one of the following occurs:*

- (1) *There exists $\mathbf{x} \geq 0$ such that $A\mathbf{x} = \mathbf{b}$,*
- (2) *There exists \mathbf{y} such that $\mathbf{y}^t \mathbf{b} < 0$ and $A^T \mathbf{y} \geq 0$.*

PROOF. Let $\mathbf{0}$ denote the zero vector in \mathbf{R}^n . Consider the primal program:

$$\text{maximize } \mathbf{0}^T \mathbf{x} \text{ subject to } A\mathbf{x} = \mathbf{b} \text{ and } \mathbf{x} \geq 0.$$

Its dual is the program:

$$\text{minimize } \mathbf{b}^T \mathbf{y} \text{ subject to } A^T \mathbf{y} \geq 0.$$

If the primal program is feasible, it is bounded and its optimal value is 0. By the generalized duality theorem (Theorem 5.2), 0 is the minimum value of $\mathbf{b}^T \mathbf{y}$ subject to $A^T \mathbf{y} \geq 0$, so (2) is ruled out. Conversely, if (2) holds, then by taking large positive multiples of \mathbf{y} we can show that $\mathbf{b}^T \mathbf{y}$ is unbounded below subject to $A^T \mathbf{y} \geq 0$. Therefore the primal program has to be infeasible. \square

DEFINITION 5.5 (Convex cone generated by a set of points). Then the convex cone generated by $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbf{R}^m$ is the set of all vectors $\mathbf{x} \in \mathbf{R}^m$ such that

$$\mathbf{x} = t_1 \mathbf{a}_1 + \dots + t_n \mathbf{a}_n \text{ for some } t_1, \dots, t_n \geq 0.$$

COROLLARY 5.6. *Let $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbf{R}^m$. Then for any point $\mathbf{x} \in \mathbf{R}^m$ exactly one of the following occurs:*

- (1) *\mathbf{x} lies in the convex cone generated by $\mathbf{a}_1, \dots, \mathbf{a}_n$.*
- (2) *There exists a hyperplane passing through the origin that separates $\mathbf{a}_1, \dots, \mathbf{a}_n$ from \mathbf{x} .*

PROOF. Take A to be the matrix with columns $\mathbf{a}_1, \dots, \mathbf{a}_n$, and apply Farkas's lemma. \square

CHAPTER 2

Flows in Networks

1. The flow maximization problem

DEFINITION 1.1 (Network). A network is a pair (N, A) , where N is set and $A \subset N^2$. We call N the set of *nodes* of the network and A the set of *arcs*. A *source* in a network is a distinguished node $u \in N$, and a target is a distinguished node $v \in N$, different from u .

The pair (N, A) is visualized by representing each node $x \in A$ by a circle labeled x , and each element of $A = (x, y)$ by an arrow joining the circle labeled x to the circle labeled y .

EXAMPLE 1.2 (Transportation networks). Let $N = \{u, 1, 2, \dots, r, 1', 2', \dots, s', v\}$, and

$$A = \{(u, i) \mid 1 \leq i \leq r\} \cup \{(i, j') \mid 1 \leq i \leq r, 2 \leq j \leq s\} \cup \{(j', v) \mid 1 \leq j \leq s\}.$$

This network, for $r = 3$ and $s = 4$, is represented by Figure 1.

DEFINITION 1.3 (Flow). A flow in a network (N, A) with source u and target v is a function $f : A \rightarrow \mathbf{R}_{\geq 0}$ whose value $f(x, y)$ at an arc $(x, y) \in A$ is known as the flow through the arc (x, y) . It is assumed to satisfy the condition that the flow

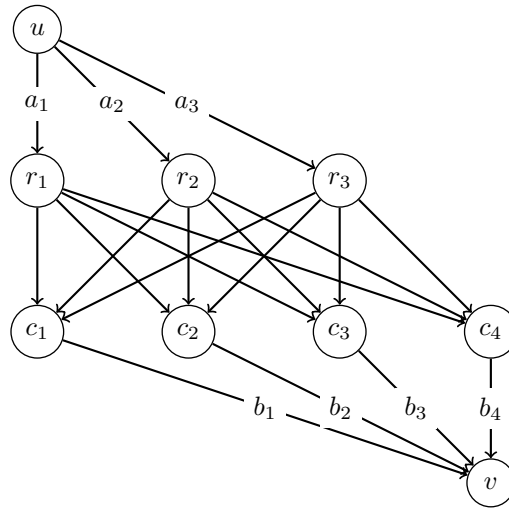


FIGURE 1. A transportation network

through any vertex other than the source or sink of the network is 0:

$$(*) \quad \sum_{\{y|(x,y) \in A\}} f(x,y) - \sum_{\{y|(y,x) \in A\}} f(y,x) = 0 \text{ for all } x \in N - \{u, v\}.$$

For subsets M_1 and M_2 of N we use the shorthand:

$$f(M_1, M_2) = \sum_{\{(x,y) \in M_1 \times M_2 | (x,y) \in A\}} f(x,y).$$

If M_1 is the singleton set $\{x\}$ we write $f(x, M_2)$ for $f(M_1, M_2)$. Similarly if $M_2 = \{y\}$ we write $f(M_1, y)$ for $f(M_1, M_2)$. Thus the equation $(*)$ can be written as:

$$f(x, N) - f(N, x) = 0 \text{ for every } x \in N.$$

The *value* of the flow is defined as the net flow out of the source u :

$$w_f = f(u, N) - f(N, u).$$

EXERCISE 1.4. Show that, for any flow f on a network (N, A) with source u and sink v , the net flow out of the source u equals the net net flow into the sink v :

$$w_f = f(u, N) - f(N, u) = f(N, v) - f(v, N).$$

DEFINITION 1.5 (Capacity function). A capacity function on a network (N, A) is a function $c : A \rightarrow (0, \infty]$ (the value ∞ is permitted). Given a capacity function, a *feasible* flow is a flow on (N, A) such that $f(x, y) \leq c(x, y)$ for every $(x, y) \in A$.

The capacity of arcs are represented by edge labels. Labels for edges with infinite capacity are omitted. In Example 1.2, set $c(u, r_i) = a_i$ for $i = 1, \dots, r$ and $c(s_j, v) = b_j$ for $j = 1, \dots, s$, and all remaining arcs to have infinite capacity. See Figure 1 for the visualization of this capacity function.

Given a network (N, A) with a capacity function $c : A \rightarrow (0, \infty]$, the maximal flow problem is to maximize w_f among all feasible flows on (N, A) with capacity c .

Thus the max-flow problem has variables $f(x, y)$, for $(x, y) \in A$, and a variable w (representing the value of the flow), and is given by:

$$(F) \quad \begin{array}{ll} f(x, N) - f(N, x) & = 0 \\ f(u, N) - f(N, u) - w & = 0, \\ f(v, N) - f(N, v) + w & = 0, \\ f(x, y) & \leq c(x, y) \end{array} \quad \begin{array}{l} \text{for every } x \in N - \{u, v\}, \\ \\ \\ \text{for every } (x, y) \in A \end{array}$$

$$\text{maximize } f(u, N) - f(N, u).$$

Let us consider the program dual to (F). It has a variable $\tau(x) \in \mathbf{R}$ for $x \in N$, and a variable $\gamma(x, y) \geq 0$ for every $(x, y) \in A$. For every (x, y) in A it has the constraint:

$$(DF) \quad \begin{array}{ll} \tau(x) - \tau(y) + \gamma(x, y) & \geq 0 \\ \tau(v) - \tau(u) & \geq 1. \end{array}$$

and objective:

$$\text{minimize } \sum_{(x,y) \in A} c(x,y)\gamma(x,y).$$

DEFINITION 1.6. Let (N, A) be a network, and $x, y \in N$. A *path* from x to y is a sequence:

$$x = x_0, x_1, \dots, x_n = y \text{ such that } (x_{i-1}, x_i) \in A \text{ for all } i = 1, \dots, n.$$

EXERCISE 1.7. Use the duality theorem to show that if there is no path from u to v in A , then the maximum value of a feasible flow is 0.

DEFINITION 1.8. A *cut* of the network (N, A) is a partition $N = X \sqcup \bar{X}$ of its nodes such that $u \in X$ and $v \in \bar{X}$. The capacity $c(X, \bar{X})$ of the cut is defined as:

$$c(X, \bar{X}) = \sum_{\{(x,y) \in A \mid x \in X, y \in \bar{X}\}} c(x, y).$$

LEMMA 1.9. If $N = X \sqcup \bar{X}$ is a cut of N , then the value of any feasible flow is bounded above by $c(X, \bar{X})$.

PROOF. Define

$$\gamma(x, y) = \begin{cases} 1 & \text{if } x \in X, y \in \bar{X}, \\ 0 & \text{otherwise.} \end{cases}$$

Define $\tau(x) = 0$ for $x \in X$ and $\tau(x) = 1$ for $x \in \bar{X}$. Then τ and γ define a feasible solution for (DF). The value of the objective is $c(X, \bar{X})$, so this must be an upper bound for (F). \square

LEMMA 1.10. Let f be a feasible flow. Define a new network (N, A_f) by:

$$(x, y) \in A_f \text{ if } \begin{cases} (x, y) \in A \text{ and } f(x, y) < c(x, y), \text{ or} \\ (y, x) \in A \text{ and } f(x, y) > 0. \end{cases}$$

If there is no path from u to v in (N, A_f) , then f is a feasible flow of maximum value.

PROOF. Let X denote the set of all nodes $x \in N$ such that (N, A_f) admits a path from u to x . The flow from X into \bar{X} is:

$$f(X, \bar{X}) = \sum_{\{(x,y) \in A \mid x \in X, y \in \bar{X}\}} f(x, y) - \sum_{\{(x,y) \in A \mid x \in \bar{X}, y \in X\}} f(x, y) = w_f.$$

Since no point from X has path to any point in \bar{X} in the network (N, A_f) , we must have:

$$\begin{aligned} f(x, y) &= c(x, y) \text{ for all } (x, y) \in A, x \in X, y \in \bar{X}, \\ f(x, y) &= 0 \text{ for all } (x, y) \in A, x \in \bar{X}, y \in X. \end{aligned}$$

Therefore $f(X, \bar{X}) = c(X, \bar{X})$ is a common value of the objective functions of (F) and its dual (DF), hence optimal for both by the duality theorem (Theorem 5.1 of Chapter 1). \square

We are now close to proving the main theorem of this section:

THEOREM 1.11 (Max flow-min cut theorem). *The maximum value for feasible flows on a network (N, A) with capacity c is equal to the minimum capacity of its cuts.*

PROOF. Let f be a feasible flow. If it satisfies the hypothesis of Lemma 1.10, then it is optimal, and result follows. If not, let $P = (u = x_0, x_1, \dots, x_n = v)$ be a path in (N, A_f) from u to v . For each arc $(x, y) \in A_f$, define its *spare capacity* as:

$$c'(x, y) = \begin{cases} c(x, y) - f(x, y) & \text{if } (x, y) \in A \text{ and } f(x, y) < c(x, y), \\ f(y, x) & \text{if } (y, x) \in A \text{ and } f(x, y) > 0. \end{cases}$$

Let $m = \min\{c'(x_{i-1}, x_i) \mid i = 1, \dots, n\}$. By construction $m > 0$. The *augmentation of f along P* is the flow \tilde{f} for which

$$\begin{aligned}\tilde{f}(x_{i-1}, x_i) &= f(x_{i-1}, x_i) + m && \text{if } (x_{i-1}, x_i) \in A \text{ and } f(x, y) < c(x, y), \\ \tilde{f}(x_i, x_{i-1}) &= f(x_i, x_{i-1}) - m && \text{if } (x_i, x_{i-1}) \in A \text{ and } f(x_i, x_{i-1}) > 0,\end{aligned}$$

for each $i = 1, \dots, n$, and all remaining values of \tilde{f} are the same as those of f . Then \tilde{f} is a feasible flow with value $w_f + m$, showing that f was not optimal. \square

The proof of the max flow-min cut theorem suggests an algorithm to find an optimal flow:

Flow Augmentation Algorithm

- (1) Begin with the zero flow $f \equiv 0$.
- (2) While the network (N, A_f) admits a path P from u to v :
 - augment f along P .
- (3) Return f .

THEOREM 1.12. *Let (N, A) be a network with capacity function c taking positive integer, or infinite values. If the set of values of feasible flows is bounded above, then the flow augmentation algorithm terminates in a finite number of steps and returns a flow with maximal value.*

PROOF. The integrality of the capacities implies that the minimum spare capacity m along a path of augmentation as in the proof of Theorem 1.11 is a positive integer. Therefore, during each run of the main loop of the algorithm the value of the resulting flow increases by 1. Therefore, the maximum possible value of all feasible flows is finite, it will be achieved in a finite number of steps. \square

REMARK 1.13. The flow augmentation algorithm returns an integer-valued flow when the capacity function is integer-valued.

Let Γ be a bipartite graph with vertex set $N = U \sqcup V$, and edge set contained in $U \times V$. Suppose $\{(x_1, y_1), \dots, (x_k, y_k)\}$ is a matching in Γ , then any vertex cover of Γ must contain at least one vertex from each of these edges. Thus the minimum size of a vertex cover of Γ must be at least as large as the maximum size of a matching. The König-Egerváry theorem states that these two quantities are equal:

THEOREM 1.14. *For a bipartite graph on vertex set $U \sqcup V$, the minimum size of a vertex cover equals the maximum number of edges in a matching in Γ .*

PROOF. Let N be the network with nodes $U \sqcup V \sqcup \{u, v\}$, with arcs (u, x) for all $x \in U$, (x, v) for all $x \in V$, and (x, y) for all $x \in U$ and $y \in V$ such that $\{x, y\}$ is an edge of Γ . Assign capacity 1 to the arcs of the form (u, x) , $x \in V$ and (x, v) , $x \in U$. Allow the arcs joining U to V infinite capacity.

A maximal value flow in (N, A) corresponds to a maximal matching in Γ (with flow of 1 along the arcs corresponding to the edges in the matching). If $X \sqcup \bar{X}$ is a cut of finite capacity, it cannot contain any arc with infinite capacity. Therefore, there are no edges in Γ between vertices from $X \cap U$ and $\bar{X} \cap V$, or in other words, every edge in Γ corresponds to an arc with both end points in X or both end points in \bar{X} . Thus $(\bar{X} \cap U) \cup (X \cap V)$ is a vertex cover. The capacity of this cut is $|\bar{X} \cap U| + |X \cap V|$. Thus the König-Egerváry theorem follows from the max flow-min cut theorem. \square

2. The supply-demand problem

Consider a network (N, A) with distinguished disjoint subsets U and V of nodes called supply nodes and demand nodes respectively. Define functions $a : U \rightarrow [0, \infty)$ and $b : V \rightarrow [0, \infty)$ representing supply and demand values at these nodes. Suppose also that a capacity function $c : A \rightarrow (0, \infty]$ is given.

The supply-demand problem is the following: determine if there exists a function $f : A \rightarrow [0, \infty)$ such that:

- (SD1) $f(u, N) - f(N, u) \leq a(u)$ for all $u \in U$,
- (SD2) $f(v, N) - f(N, v) \leq -b(v)$ for all $v \in V$,
- (SD3) $f(x, N) - f(N, x) = 0$ for all $x \in N - (U \cup V)$,
- (SD4) $f(x, y) \leq c(x, y)$ for all $(x, y) \in A$.

This is a feasibility problem, and so we take the objective function to be identically 0. The dual problem has variables $\tau(x) \in \mathbf{R}$ for $x \in N$ and $\gamma(x, y) \geq 0$ for $(x, y) \in A$, satisfying:

$$\tau(x) - \tau(y) + \gamma(x, y) \geq 0 \text{ for all } (x, y) \in A.$$

The objective is:

$$\text{minimize } \sum_{(x,y) \in A} c(x,y)\gamma(x,y) + \sum_{u \in U} a(u)\tau(u) - \sum_{v \in V} b(v)\tau(v).$$

Suppose that $N = X \cup \bar{X}$ is any partition of N into two parts. As usual, let $\gamma(x, y) = 1$ if $x \in X$ and $y \in \bar{X}$, and 0 otherwise. Let $\tau(x) = 0$ for $x \in X$ and $\tau(x) = 1$ for $x \in \bar{X}$. These are feasible values for the dual program, and the value of the objective function is:

$$a(\bar{X} \cap U) - b(\bar{X} \cap V) + c(X, \bar{X}).$$

If the primal program were feasible, then the maximum value of its objective function, namely 0, would be a lower bound for the objective function of the dual. Therefore, if the supply demand problem has a feasible solution then:

$$(10) \quad b(\bar{X} \cap V) - a(\bar{X} \cap U) \leq c(X, \bar{X}) \text{ for every cut } N = X \cup \bar{X}.$$

In other words, for every subset \bar{X} of N , the shortfall in supply from nodes in \bar{X} against demand from nodes in \bar{X} cannot exceed the amount that can be supplied across the cut $X \rightarrow \bar{X}$.

THEOREM 2.1 (Gale's supply-demand theorem). *The supply-demand problem has a feasible solution if and only if the condition (10) is satisfied.*

PROOF. Enlarge the network (N, A) to (N^*, A^*) , where $N^* = N \cup \{u_0, v_0\}$, and $A^* = A \cup \{(u_0, u) \mid u \in U\} \cup \{(v, v_0) \mid v \in V\}$. For any cut $N^* = X^* \cup \bar{X}^*$, let $X = X^* \cap N$ and $\bar{X} = \bar{X}^* \cap N$.

$$\begin{aligned} c(X^*, \bar{X}^*) &= a(\bar{X} \cap U) + b(X \cap V) + c(X, \bar{X}) \\ &= a(\bar{X} \cap U) - b(\bar{X} \cap V) + c(X, \bar{X}) + b(V) \\ &\geq b(V) = c(\{u_0\} \cup N, \{v_0\}), \end{aligned}$$

the last line being the capacity of the cut $(\{u_0\} \cup N, \{v_0\})$ separating v_0 from all the nodes in N^* . Therefore $(\{u_0\} \cup N, \{v_0\})$ is a cut in (N^*, A^*) of minimum capacity $b(V)$. It follows that (N^*, A^*) has a feasible flow with value $b(V)$. The restriction of this flow to (N, A) is a feasible solution to the supply-demand problem. \square

REMARK 2.2. If the functions a and b are integer-valued, then the supply-demand problem has an integer-valued feasible solution.

COROLLARY 2.3. *Given vectors $\mathbf{a} \in \mathbf{R}^r$, $\mathbf{b} \in \mathbf{R}^s$ with positive coordinates having the same sum, let $\mathbf{N}_{\mathbf{ab}}$ denote the set of all $r \times s$ matrices x_{ij} such that*

$$\begin{aligned} 0 \leq x_{ij} \leq 1 \text{ for all } 1 \leq i \leq r, 1 \leq j \leq s, \\ \sum_{j=1}^s x_{ij} = a_i \text{ for all } 1 \leq i \leq r, \\ \sum_{i=1}^r x_{ij} = b_j \text{ for all } 1 \leq j \leq s. \end{aligned}$$

Then $\mathbf{N}_{\mathbf{ab}}$ is non-empty if and only if, for every subset $R \subset [r]$ and every subset $S \subset [s]$,

$$\sum_{j \in S} b_j - \sum_{i \in R} a_i \leq (r - |R|)|S|.$$

Moreover, $\mathbf{N}_{\mathbf{ab}}$ is non-empty if and only if it has an integer point.

PROOF. Consider a network (N, A) whose set N of nodes is the disjoint union of $[r]$ and $[s]$. Let $U = [r]$ and $V = [s]$. Let $A = \{(i, j) \mid i \in [r], j \in [s]\}$. Define $c(i, j) = 1$, $a(i) = a_i$, $b(j) = b_j$ for $i \in [r]$ and $j \in [s]$. Let f be a feasible solution to this supply demand problem. Setting $x_{ij} = f(i, j)$ gives a point in $\mathbf{N}_{\mathbf{ab}}$.

By Theorem 2.1, this exists if and only if, for every subset $X \subset [r] \cup [s]$,

$$b(\bar{X} \cap [s]) - a(\bar{X} \cap [r]) \leq c(X, \bar{X}).$$

setting $R = \bar{X} \cap [r]$ and $S = \bar{X} \cap [s]$ gives the condition in the theorem. \square

COROLLARY 2.4. *There exists a function $f : A \rightarrow [0, \infty)$ satisfying (SD1)–(SD4) if and only if, for every subset $V' \subset V$, there exists a function $f_{V'} : A \rightarrow [0, \infty)$ satisfying (SD1), (SD3), (SD4) such that*

$$(SD5) \quad f_{V'}(N, V') - f_{V'}(V', N) \geq b(V').$$

REMARK 2.5. The inequality (SD5) says that, for every subset V' of demand nodes, there exists a function satisfying the supply conditions, the capacity conditions, and the “flow condition” (SD3) which meets the aggregate demand of V' (but not necessarily the node-by-node demands).

PROOF. If f satisfies (SD1)–(SD4), then it obviously satisfies (SD5).

Conversely, suppose such an $f_{V'}$ exists for every $V' \subset V$. Let $N = X \cup \bar{X}$ be a partition. Take $V' = V \cap \bar{X}$, and let $N_0 = N - (U \cup V)$. Then, using (SD1), (SD3), and (SD5) we have:

$$\begin{aligned} -a(U \cap \bar{X}) &\leq f_{V'}(N, U \cap \bar{X}) - f_{V'}(U \cap \bar{X}, N), \\ 0 &= f_{V'}(N, N_0 \cap \bar{X}) - f_{V'}(N_0 \cap \bar{X}, N), \\ b(V \cap \bar{X}) &\leq f_{V'}(N, V \cap \bar{X}) - f_{V'}(V \cap \bar{X}, N). \end{aligned}$$

Adding up gives:

$$b(V \cap \bar{X}) - a(U \cap \bar{X}) \leq f_{V'}(N, \bar{X}) - f_{V'}(\bar{X}, N) = f_{V'}(X, \bar{X}) - f_{V'}(\bar{X}, X).$$

Finally (SD4) implies that $f_{V'}(X, \bar{X}) - f_{V'}(\bar{X}, X) \leq c(X, \bar{X})$. Thus the inequality (10) is satisfied and the corollary follows from Theorem 2.1. \square

A dual version of Corollary 2.4 is

COROLLARY 2.6. *There exists a function $f : A \rightarrow [0, \infty)$ satisfying (SD1)–(SD4) if and only if, for every subset $U' \subset U$, there exists a function $f_{U'} : A \rightarrow [0, \infty)$ satisfying (SD2), (SD3), (SD4) such that*

$$(SD5') \quad f_{U'}(U', N) - f_{U'}(N, U') \leq a(U').$$

EXERCISE 2.7. Prove Corollary 2.6.

EXAMPLE 2.8 (Hall's Marriage Theorem). Let $U = \{1, \dots, r\}$, $V = \{1', \dots, r'\}$. Let Γ be a bipartite graph with vertex set $U \sqcup V$, Hall's marriage theorem says that Γ contains a perfect matching if and only if, for every $V' \subset V$, there exists $U' \subset U$ such that $|U'| = |V'|$ and every vertex in V is connected to at least one vertex in U' .

Hall's marriage theorem is a special case of Corollary 2.1. To see this, let $N = U \sqcup V$, $A = \{(i, j) \mid i \text{ is connected to } j \text{ in } \Gamma\}$. Set the capacity of each arc to be 1. Let $\mathbf{a} = \mathbf{b} = (1, \dots, 1)$, the all-ones vector. Then an integer solution to the supply demand problem is a perfect matching. Suppose, for every subset $V' \subset V$, there exists $U' \subset U$ such that $|U'| = |V'|$ and for every $x \in U'$, there exists $y \in V'$ such that $(x, y) \in A$. Set $f_{V'}(x, y) = 1$ for each such arc. Then $f_{V'}$ satisfies the conditions in Corollary 2.4, so that there exists a feasible solution to the supply demand problem.

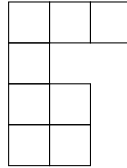
Given a vector $\mathbf{a} \in \mathbf{R}^r$, let \mathbf{a}^\downarrow denote the vector obtained from \mathbf{a} by replacing its entries in weakly decreasing order. Let a_i^\downarrow denote the i th coordinate of \mathbf{a}^\downarrow (the i th order statistic of \mathbf{a}).

DEFINITION 2.9. Given two vectors $\mathbf{a} \in \mathbf{R}^r$ and $\mathbf{b} \in \mathbf{R}^s$, we say that \mathbf{a} majorizes \mathbf{b} if, \mathbf{a} and \mathbf{b} have the same sum, and for every $1 \leq k \leq r$,

$$\sum_{i=1}^k a_i^\downarrow \geq \sum_{i=1}^k b_i^\downarrow.$$

We write $\mathbf{a} \triangleright \mathbf{b}$.

A visual representation of a non-negative integer vector $\mathbf{a} \in \mathbf{R}^r$ is by its *Young diagram*. This is a configuration of r left-justified rows of boxes, the i th row consisting of a_i boxes. For example, if $\mathbf{a} = (3, 1, 2, 2)$, then the Young diagram is:



DEFINITION 2.10 (Conjugate of an integer vector). Given $\mathbf{a} \in \mathbf{R}^r$, Define $\mathbf{a}' = (a'_1, a'_2, \dots)$ by $a'_1 + \dots + a'_k = \sum_{i=1}^r \min(a_i, k)$. Thus a'_k is the number of boxes in the k th column of the Young diagram of \mathbf{a} . Then \mathbf{a}' is an integer partition, known as the *conjugate* of \mathbf{a} .

For example, the conjugate of $(3, 1, 2, 2)$ is $(4, 3, 1)$. Note that rearranging the coordinates of \mathbf{a} does not change its conjugate. This the conjugate of $(3, 2, 1, 1)$ is also $(4, 3, 1)$.

EXERCISE 2.11. If \mathbf{a} is an integer partition, then show that $(\mathbf{a}')' = \mathbf{a}$.

THEOREM 2.12 (Gale-Ryser). *The polytope $N_{\mathbf{a}\mathbf{b}}$ is non-empty if and only if $\mathbf{a}' \triangleright \mathbf{b}$.*

Suppose $V' \subset V$. A flow $f_{V'}$ which satisfies (SD1), (SD3), and (SD4) of maximum value is obtained by simply supplying as much as possible from each a_i to V' . If $|V'| = k$, then $f_{V'}(u_i, V') = \min(a_i, k)$. Summing over all $u \in U$ gives $f_{V'}(N, V') = f_{V'}(U, V') = \sum_{i=1}^r \min(a_i, k)$. The condition (SD5) becomes:

$$(11) \quad \sum_{i=1}^k \min(a_i, k) \geq \sum_{j \in V'} b_j \text{ for every } V' \subset V \text{ with } |V'| = k.$$

Then the left hand side of (11) is the number of boxes in the first r columns of the Young diagram of a . The condition (11) becomes:

$$\sum_{i=1}^k a'_i \geq \sum_{j \in V'} b_j \text{ for every } V' \subset V \text{ with } |V'| = k.$$

Then the above inequalities are equivalent to:

$$\sum_{i=1}^k a'_i \geq \sum_{i=1}^k b_i^\downarrow.$$

In other words, the partition $\mathbf{a}' \triangleright \mathbf{b}$.

EXERCISE 2.13. Let $\mathbf{a} \in \mathbf{R}^r$ and $\mathbf{b} \in \mathbf{R}^s$ be non-negative integer vectors with equal sum d . Reinterpret the Gale-Ryser theorem in the following forms:

- (1) There exist set partitions $[d] = R_1 \sqcup \cdots \sqcup R_r = S_1 \sqcup \cdots \sqcup S_s$ such that $|R_i| = a_i$ for each i , and $|S_j| = b_j$ for each j , and $|A_i \cap B_j| \leq 1$ if and only if $\mathbf{a}' \triangleright \mathbf{b}$.
- (2) There exists a bipartite simple graph on vertex set $[r] \sqcup [s]$ such that the degree of each $i \in [r]$ is a_i and the degree of $j \in [s]$ is b_j if and only if $\mathbf{a}' \triangleright \mathbf{b}$.

3. The Hitchcock Problem

Given vectors $\mathbf{a} \in \mathbf{Z}_{\geq 0}^r$ and $\mathbf{b} \in \mathbf{Z}_{\geq 0}^s$ with the same sum, and a *cost array* $(a_{ij})_{r \times s}$, the Hitchcock problem is the following linear program in variables $f_{ij} \geq 0$:

$$(H) \quad \begin{array}{ll} -\sum_{j=1}^s f_{ij} & \geq -a_i \text{ for } i = 1, \dots, r, \\ \sum_{i=1}^r f_{ij} & \geq b_j \text{ for } j = 1, \dots, s, \\ \text{minimize} & \sum_{i=1}^r \sum_{j=1}^s a_{ij} f_{ij}. \end{array}$$

The constraints are like in the supply demand problems in Section 2, but now there is an objective function with positive coefficients.

The dual program has variables $\alpha_i > 0$ for $i = 1, \dots, r$ and $\beta_j \geq 0$ for $j = 1, \dots, s$, and is given by:

$$(DH) \quad \begin{array}{ll} -\alpha_i + \beta_j & \geq a_{ij}, \\ \text{maximize} & -\sum_{i=1}^r \alpha_i a_i + \sum_{j=1}^s b_j \beta_j. \end{array}$$

Suppose $(\alpha_1, \dots, \alpha_r, \beta_1, \dots, \beta_s)$ is a feasible solution for (DH), and (f_{ij}) is a feasible solution for (H) such that:

$$(12) \quad -\alpha_i + \beta_j < a_{ij} \implies f_{ij} = 0.$$

Then,

$$\begin{aligned}
\sum_{i=1}^r \sum_{j=1}^s a_{ij} f_{ij} &= \sum_{-\alpha_i + \beta_j < a_{ij}} a_{ij} f_{ij} + \sum_{-\alpha_i + \beta_j = a_{ij}} a_{ij} f_{ij} \\
&= 0 + \sum_{-\alpha_i + \beta_j = a_{ij}} (-\alpha_i + \beta_j) f_{ij} \\
&= -\sum_{i=1}^r \alpha_i a_i + \sum_{j=1}^s \beta_j b_j.
\end{aligned}$$

Thus we have feasible points for (H) and (DH) where the objective functions coincide, hence are optimal.

A strategy to solve the Hitchcock problem is to try and find such a pair of solutions. Begin with an obvious feasible solution to (DH): Take $\alpha_i = 0$ for all i , and $\beta_j = \min\{a_{ij} \mid i = 1, \dots, r\}$ for $j = 1, \dots, s$. Now use the flow augmentation algorithm on the network N^* from the proof of Theorem 2.1 to try and find a feasible solution with $f_{ij} \geq 0$ to:

$$\begin{aligned}
(13) \quad & -\sum_{j=1}^s f_{ij} \geq -a_i \text{ for } i = 1, \dots, r, \\
& \sum_{i=1}^r f_{ij} \geq b_j \text{ for } j = 1, \dots, s, \\
& f_{ij} = 0 \text{ if } -\alpha_i + \beta_j = a_{ij}.
\end{aligned}$$

This is a supply demand problem in a modified transportation network (Example 1.2) where some of the arcs (i, j') are blocked.

If such a solution exists, then, as explained above, we have feasible solutions to (H) and (DH) with equal values of their objective functions, and we are done.

Otherwise, proceeding as in the proof of Lemma 1.10, Let $X \subset N^*$ denote the set of nodes reachable from u_0 in the modified transportation network where arcs (i, j') are blocked when $-\alpha_i + \beta_j < a_{ij}$. Increase α_i and β_j by 1 for all the nodes in \bar{X} . Since all the arcs between X and \bar{X} had slack ($-\alpha_i + \beta_j < a_{ij}$ when $i \in X$ and $j' \in \bar{X}$), the modified values of the dual variable are still feasible.

Now go back and to solving (13) with the modified dual variables. Continue this process until either an optimal common value for the objective functions of (H) and (DH) are found. If this algorithm fails to produce a flow even when $-\alpha_i + \beta_j = a_{ij}$ for all i, j , then the transportation problem itself is infeasible.

4. The cost minimization problem

The Hitchcock problem is a special case of a more general minimal cost supply-demand problem, where the constraints (SD1)–(SD4) of Section 2 on variables $f(x, y) \geq 0$ are paired with the objective:

$$\text{maximize } \sum_{(x,y) \in A} a(x, y) f(x, y).$$

Assume that the cost function $a : A \rightarrow [0, \infty)$ and the capacity function $c : A \rightarrow (0, \infty]$ are integer valued (but c is allowed infinite values). Also the supply and demand constraints a and b are integer-valued.

We modify the problem a little bit for convenience. Add nodes u_0 and v_0 to the network, and connect u_0 to each node $u \in U$ by an arc of capacity $a(u)$ and cost 0. Connect v_0 to each node $v \in V$ with an arc of capacity $b(v)$ and cost 0. Denote the resulting network by (N^*, A^*) . We introduce an auxiliary variable w

for the value of the flow. The modified program has variables $f(x, y) \geq 0$ for each $(x, y) \in A^*$, and a variable $w \geq 0$, and is given by:

$$(C) \quad \begin{array}{rcl} f(u, N) - f(N, u) & = & v, \\ f(x, N) - f(N, x) & = & 0 \text{ for } x \in N^* - \{u, v\}, \\ f(v, N) - f(N, v) & = & v, \\ f(x, y) & \leq & c(x, y) \text{ for } (x, y) \in A^*. \end{array} \quad \frac{\text{minimize} \quad pw - \sum_{(x, y) \in A} a(x, y) f(x, y).}{}$$

The quantity p is a parameter, set to take values $0, 1, 2, \dots$ sequentially. As p increases, minimal cost flows of successively higher values are obtained, and for p sufficiently large, flows of maximal value are obtained (without reference to the cost function). The dual has variables $\tau(x) \in \mathbf{R}$ for $x \in N^*$, and $\gamma(x, y) \geq 0$ for $(x, y) \in A^*$ and is given by:

$$(DC) \quad \begin{array}{rcl} -\tau(u) + \tau(v) & = & p, \\ \tau(x) - \tau(y) + \gamma(x, y) & \geq & -a(x, y) \text{ for } (x, y) \in A^*, \end{array} \quad \frac{\text{minimize} \quad \sum_{(x, y) \in A^*} c(x, y) \gamma(x, y).}{}$$

Bibliography

- [1] B. Gärtner and J. Matousek. *Understanding and Using Linear Programming*. Universitext. Springer-Verlag, Berlin, 2007.
- [2] J. A. D. Loera and E. D. Kim. Combinatorics and geometry of transportation polytopes: An update. In A. Barg and O. R. Musin, editors, *Discrete Geometry and Algebraic Combinatorics*, volume 625 of *Contemporary Mathematics*, pages 37–76. American Mathematical Society, 2014.
- [3] R. A. Brualdi. *Combinatorial Matrix Classes*. Encyclopedia of Mathematics and its Applications, 108. Cambridge University Press, Cambridge, 2006.
- [4] L. R. Ford, Jr., and D. R. Fulkerson *Flows in Networks*. Princeton Landmarks in Mathematics. Princeton University Press, Princeton, 2010.