2 Total Unimodularity (TU) and Its Applications

In this section we will discuss the total unimodularity theory and its applications to flows in networks.

2.1 Total Unimodularity: Definition and Properties

Consider the following integer linear programming problem

$$(P) \quad \begin{array}{ll} \max & c^T x \\ \text{s.t.} & Ax = b \\ & x \geq 0 \end{array} \tag{2.1}$$

where $A \in Z^{m\times n}$, $b \in Z^m$ and $C \in Z^n$ all integers.

Definition 2.1 *A square, integer matrix B is called* **unimodular** *if $|\mathrm{Det}(B)| = 1$. An integer matrix A is called* **totally unimodular** *if every square, nonsingular submatrix of A is unimodular.*

The above definition means that a TU matrix is a $\{1, 0, -1\}$-matrix. But, a $\{1, 0, -1\}$-matrix may not necessarily a TU matrix, e.g.,

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

Lemma 2.1 *Suppose that $A \in Z^{n\times n}$ is a unimodular matrix and that $b \in Z^n$ is an integer vector. If A is nonsingular, then $Ax = b$ has the unique integer solution $x = A^{-1}b$.*

Proof. Let a_{ij} be the ij-th entry of A, $i, j = 1, \ldots, n$. For any a_{ij}, define the cofactor of a_{ij} as

$$\mathrm{Cof}(a_{ij}) = (-1)^{i+j}\mathrm{Det}(A^{\{1,\ldots,n\}\setminus\{j\}}_{\{1,\ldots,n\}\setminus\{i\}}),$$

where $(A^{\{1,\ldots,n\}\setminus\{j\}}_{\{1,\ldots,n\}\setminus\{i\}})$ is the matrix obtained by removing the i-th row and the j-th column of A. Then

$$\mathrm{Det}(A) = \sum_{i=1}^{n} a_{i1} \times \mathrm{Cof}(a_{i1}).$$

The Adjoint of A is

$$\mathrm{Adj}(A) = \mathrm{Adj}(\{a_{ij}\}) = \{\mathrm{Cof}(a_{ij})\}^T$$

and the inverse of A is

$$A^{-1} = \frac{1}{\mathrm{Det}(A)}\mathrm{Adj}(A).$$

Since $A \in Z^{n\times n}$ is a unimodular nonsingular integer matrix, every $\mathrm{Cof}(a_{ij})$ is an integer and $\mathrm{Det}(A) = \pm 1$. Hence A^{-1} is an integer matrix and $x = A^{-1}b$ is integer whenever b is. Q.E.D.

Theorem 2.1 *If A is TU, every basic solution to P is integer.*

Proof. Suppose that x is a basic solution to P. Let N be the set of indices of x such that $x_j = 0$. Since x is a basic solution to P, there exist two nonnegative integers p and q with $p+q = n$ and indices $B(1), \ldots, B(p) \in \{1, \ldots, m\}$ and $N(1), \ldots, N(q) \in N$ such that

$$\{A^T_{B(i)}\}_{i=1}^p \cup \{e^T_{N(j)}\}_{j=1}^q$$

are linearly independent, where $e_{N(j)}$ is the $N(j)$-th unit vector in $\Re^n$.

From $Ax = b$ and $x_j = 0$ for $j \in N$, we know that there exists a matrix $B \in \Re^{p\times p}$ such that

$$Bx_B = b_B,$$

where $x_B = (x_{B(1)}, \ldots, x_{B(p)})^T$ and $b_B = (b_{B(1)}, \ldots, b_{B(p)})^T$. The matrix B is nonsingular from the linear independence of $\{A^T_{B(i)}\}_{i=1}^p \cup \{e^T_{N(j)}\}_{j=1}^q$. Then, by Lemma 2.1, we know that x_B is integer. By noting that $x_N = 0$ is integer, we complete the proof. Q.E.D.

Proposition 2.1 *$A \in Z^{m\times n}$ is TU $\Longrightarrow -A$ and A^T are totally unimodular matrices.*

Proposition 2.2 *$A \in Z^{m\times n}$ is TU $\Longrightarrow (A\ e_i)$ is TU, where e_i is the i-th unit vector of $\Re^m$, $i = 1, \ldots, m$.*

Proposition 2.3 *$A \in Z^{m \times n}$ is TU $\Longrightarrow$ $(A\ I)$ is TU, where $I \in \Re^{m \times m}$ is the identity matrix.*

Proposition 2.4 *$A \in Z^{m \times n}$ is TU $\Longrightarrow$ $\begin{pmatrix} A \\ I \end{pmatrix}$ is TU, where $I \in \Re^{n \times n}$ is the identity matrix.*

Theorem 2.2 *(Hoffman and Kruskal, 1956) For any integer matrix $A \in Z^{m \times n}$, the following statements are equivalent:*

1. *A is TU;*
2. *The extreme points (if any) of $S(b) = \{x|\ Ax \leq b,\ x \geq 0\}$ are integer for any integer b;*
3. *Every square nonsingular submatrix of A has integer inverse.*

Proof.
(1 $\Rightarrow$ 2)

After adding nonnegative slack variables, we have the system

$$Ax + Is = b,\ x \geq 0,\ s \geq 0.$$

The extreme points of $S(b)$ correspond to basic feasible solutions of the system (as an exercise). Let $y = (x, s)$ be a basic feasible solution of the above system. If a given basis B contains only columns from A, then y_B is integer as A is TU (Lemma 2.1). The same is true if B contains only columns from I. So we have to consider the case when $B = (\bar{A}\ \bar{I})$, where $\bar{A}$ is a submatrix of A and $\bar{I}$ is a submatrix of I. After the permutation of rows of B, we have

$$B' = \begin{pmatrix} A_1 & 0 \\ A_2 & I' \end{pmatrix}.$$

Obviously, $|\text{Det}(B)| = |\text{Det}(B')|$ and

$$|\text{Det}(B')| = |\text{Det}(A_1)||\text{Det}(I')| = |\text{Det}(A_1)|.$$

Now A is totally unimodular implies $|\text{Det}(A_1)| = 0$ or 1 and since B is assumed to be nonsingular, $|\text{Det}(B')| = 1$. Again, from Lemma 2.1, y_B is an integer. Hence y is integer because $y_j = 0, j \notin B$. This implies that x is integer. [One may also make use of Theorem 2.1 and Proposition 2.3 to get the proof immediately.]

(**2** $\Rightarrow$ **3**).

Let $B \in Z^{p \times p}$ be any square nonsingular submatrix of A. It is sufficient to prove that $\bar{b}_j$ is an integer vector, where $\bar{b}_j$ is the jth column of B^{-1}, $j = 1, \ldots, p$.

Let t be an integer vector such that $t + \bar{b}_j > 0$ and $b_B(t) = Bt + e_j$, where e_j is the jth unit vector. Then

$$x_B = B^{-1}b_B(t) = B^{-1}(Bt + e_j) = t + B^{-1}e_j = t + \bar{b}_j > 0.$$

By choosing b_N ($N = \{1, \ldots, n\} \backslash B$) sufficiently large such that $(Ax)_j < b_j$, $j \in N$, where $x_j = 0$, $j \in N$. Hence x is an extreme point of $S(b(t))$. As x_B and t are integer vectors, $\bar{b}_j$ is an integer vector too for $j = 1, \ldots, p$ and B^{-1} is an integer.

(**3** $\Rightarrow$ **1**).

Let B be an arbitrary square, nonsingular submatrix of A. Then

$$1 = |\text{Det}(I)| = |\text{Det}(BB^{-1})| = |\text{Det}(B)||\text{Det}(B^{-1})|.$$

By the assumption, B and B^{-1} are integer matrices. Thus

$$|\text{Det}(B)| = |\text{Det}(B^{-1})| = 1,$$

and A is TU. Q.E.D.

Theorem 2.3 *(A sufficient condition of TU) An integer matrix A with all $a_{ij} = 0, 1,$ or -1 is TU if*

1. *no more than two nonzero elements appear in each column,*
2. *the rows of A can be partitioned into two subsets M_1 and M_2 such that*
   (a) *if a column contains two nonzero elements with the same sign, one element is in each of the subsets,*
   (b) *if a column contains two nonzero elements of opposite signs, both elements are in the same subset.*

Proof. The proof is by induction. One element submatrix of A has a determinant equal to $(0, 1, -1)$.

Assume that the theorem is true for all submatrices of A of order $k-1$ or less. If B contains a column with only one nonzero element, we expand $\mathrm{Det}(B)$ by that column and apply the induction hypothesis.

Finally, consider the case in which every column of B contains two nonzero elements. Then from 2(a) and 2(b) for every column j

$$\sum_{i\in M_1} b_{ij} = \sum_{i\in M_2} b_{ij}, \quad j = 1, \ldots, k.$$

Let b_i be the ith row. Then the above equality gives

$$\sum_{i\in M_1} b_i - \sum_{i\in M_2} b_i = 0,$$

which implies that $\{b_i\}$, $i \in M_1 \cup M_2$ are linearly dependent and thus B is singular, i.e., $\mathrm{Det}(B) = 0$. Q.E.D.

Corollary 2.1 *The vertex-edge incidence matrix of a bipartite graph is TU.*

Corollary 2.2 *The node-arc incidence matrix of a digraph is TU.*

2.2 Applications

In this section we show that the assumptions in Theorems in Section 2.1 for integer programming problems connected with optimization of flows in networks are fulfilled. This means that these problems can be solved by the **SIMPLEX METHOD**. However, it is not necessarily to use the simplex method because more efficient methods have been developed by taking into consideration the specific structure of these problems.

Many commodities, such as gas, oil, etc., are transported through networks in which we distinguish sources, intermediate transportation or distribution points and destination points.

We will represent a network as a directed graph $G = (V, E)$ and associate with each arc $(i, j) \in E$ the flow x_{ij} of the commodity and the capacity d_{ij} (possibly infinite) that bounds the flow through the arc. The set V is partitioned into three sets:

- V_1 −− set of sources or origins,
- V_2 −− set of intermediate points,
- V_3 −− set of destinations or sinks.

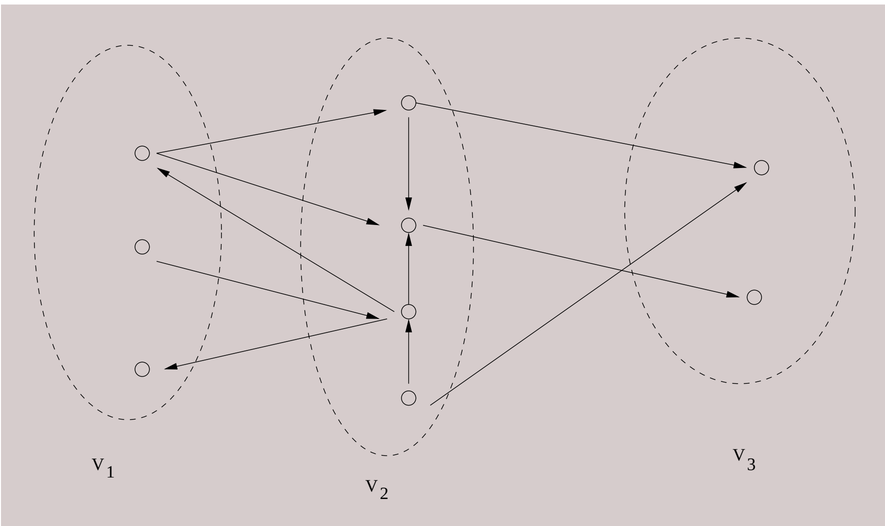

Figure 2.1: A network

For each $i \in V_1$, let a_i be a supply of the commodity and for each $i \in V_3$, let b_i be a demand for the commodity.
We assume that there is no loss of the flow at intermediate points. Additionally, denote $V(i)$ $(V'(i))$ as

$$V(i) = \{j|\ (i,j) \in E\} \quad \text{and} \quad V'(i) = \{j|\ (j,i) \in E\},$$

respectively.

Then the **minimum cost capacitated problem** may be formulated as

$$\text{(P)} \qquad v(P) = \min \sum_{(i,j)\in E} c_{ij} x_{ij}$$

subject to

$$\sum_{j\in V(i)} x_{ij} - \sum_{j\in V'(i)} x_{ji} \begin{cases} \leq a_i, & i \in V_1, \\ = 0, & i \in V_2, \\ \leq -b_i, & i \in V_3, \end{cases} \tag{2.2}$$

$$0 \leq x_{ij} \leq d_{ij}, \quad (i,j) \in E. \tag{2.3}$$

Constraint (2.2) requires the conservation of flow at intermediate points, a net flow into sinks at least as great as demanded, and a net flow out of sources equal or less than the supply. In some applications, demand must be satisfied exactly and all of the supply must be used. If all of the constraints of (2.2) are equalities, the problem has no feasible solutions unless

$$\sum_{i\in V_1} a_i = \sum_{i\in V_3} b_i.$$

To avoid pathological cases, we assume for each cycle in the network $G = (V, E)$ either that the sum of costs of arcs in the cycle is positive or that the minimal capacity of an arc in the cycle is bounded.

Theorem 2.4 *The constraint matrix corresponding to (2.2) and (2.3) is totally unimodular.*

Proof. The constraint matrix has the form

$$A = \begin{bmatrix} A_1 \\ I \end{bmatrix},$$

where A_1 is the matrix for (2.2) and I is an identity matrix for (2.3). In the last section, we show that A_1 is totally unimodular implies that A is totally unimodular.

Each variable x_{ij} appears in exactly two constraints of (2.2) with coefficients $+1$ or -1. Thus $-A_1$ is an incidence matrix for a digraph and therefore it is totally unimodular. Q.E.D.

The most popular case of P is the so-called (capacitated) **transportation problem.** We obtain it if we put in P: $V_2 = \emptyset$, $V'(i) = \emptyset$ for all $i \in V_1$ and $V(i) = \emptyset$ for all $i \in V_3$.

So we get

$$\text{(TP)} \qquad \begin{aligned} v(T) = \min \quad & \sum_{(i,j)\in E} c_{ij}x_{ij}, \\ \text{s.t.} \quad & \sum_{j\in V(i)} x_{ij} \le a_i, \;\; i \in V_1, \\ & \sum_{j\in V'(i)} x_{ji} \ge b_i, \;\; i \in V_3, \\ & 0 \le x_{ij} \le d_{ij}, \;\; (i,j) \in E. \end{aligned}$$

If $d_{ij} = \infty$ for all $(i,j) \in E$, the uncapacitated version of P is sometimes called the **transshipment problem**.

If all $a_i = 1$, and all $b_i = 1$, and additionally, $|V_1| = |V_3|$, the transshipment problem reduces to the so-called **assignment problem** of the form

$$
\text{(AP)}\qquad
\begin{aligned}
v(AP) = \min\ & \sum_{i \in V_1} \sum_{j \in V(i)} c_{ij} x_{ij}, \\
\text{s.t.}\ & \sum_{j \in V(i)} x_{ij} = 1,\ i \in V_1, \\
& \sum_{j \in V'(i)} x_{ji} = 1,\ i \in V_3, \\
& x_{ij} \geq 0.
\end{aligned}
$$

Note that $|V_1| = |V_3|$ implies that all constraints in (AP) must be satisfied as equalities.

Let $V = \{1, \ldots, m\}$. Still another important practical problem obtained from P is called the **maximum flow problem**. In this problem, $V_1 = \{1\}$, $V_3 = \{m\}$, $V'(1) = \emptyset$, $V(m) = \emptyset, a_1 = \infty$, $b_m = \infty$.

The problem is to maximize the total flow into the vertex m under the capacity constraints

$$
\text{(MF)}\qquad
\begin{aligned}
v(MF) = \max & \sum_{i \in V'(m)} x_{im}, \\
\text{s.t.} & \sum_{j \in V(i)} x_{ij} - \sum_{j \in V'(i)} x_{ji} = 0, \\
& i \in V_2 = \{2, \ldots, m-1\}, \\
& 0 \leq x_{ij} \leq d_{ij},\ (i,j) \in E.
\end{aligned}
$$

Finally, consider the **shortest path problem**. Let c_{ij} be interpreted as the length of edge (i,j). Define the length of a path in G to be the sum of the edge lengths over all edges in the path. The objective is to find a path of minimum length from a vertex 1 to vertex m. It is assumed that all cycles have nonnegative length. This problem is a special case of the transshipment problem in which $V_1 = \{1\}$, $V_3 = \{m\}$, $a_1 = 1$ and $b_m = 1$.

Let A be the incidence matrix of the digraph $G = (V, E)$, where $V = \{1, \ldots, m\}$ and $E = \{e_1, \ldots, e_n\}$. With each arc e_j we associate its length $c_j \geq 0$ and its flow $x_j \geq 0$. The shortest path problem may be formulated as:

$$v(SP) = \min \sum_{j=1}^{n} c_j x_j,$$

$$\text{(SP)} \qquad \text{s.t.} \qquad Ax = \begin{bmatrix} -1 \\ 0 \\ \vdots \\ 0 \\ +1 \end{bmatrix}, \; x \geq 0.$$

The first constraint corresponds to the source vertex, the mth constraint corresponds to the demand vertex, while the remaining constraints correspond to the intermediate vertices, i.e., the points of distribution of the unit flow.

The dual problem to SP is

$$\text{(DSP)} \qquad v(DSP) = \max(-u_1 + u_m),$$

$$A^T u \leq c. \tag{2.4}$$