

ON SOME CHARACTERISATIONS OF TOTALLY UNIMODULAR MATRICES

D. de WERRA

Département de Mathématiques, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

Received 30 August 1978

Revised manuscript received 22 April 1980

A characterisation of totally unimodular matrices is derived from a result of Hoffman and Kruskal. It is similar in spirit to a result of Baum and Trotter. Its relation with some other known characterisations is discussed and in the particular case where the matrices have $(0, 1)$ entries, we derive some properties of the associated unimodular hypergraphs. Similar results for balanced and perfect matrices are also reviewed.

Key words: Totally Unimodular Matrices, Balanced Matrices, Hypergraphs, Perfect Graphs, Coloring.

1. Introduction

The purpose of this note is to review and establish some characterisations of totally unimodular (t.u.) matrices. The main tool for this will be the theorem of Hoffman and Kruskal on t.u. matrices.

Here $\lfloor t \rfloor$ (resp. $\lceil t \rceil$) will denote the lower (resp. upper) integer part of real number t . If c is a vector of real numbers c_i , then $\lfloor c \rfloor$ (resp. $\lceil c \rceil$) will denote the vector whose components are $\lfloor c_i \rfloor$ (resp. $\lceil c_i \rceil$).

Recall that a matrix A is t.u. if the determinant of every square submatrix of A has value 0, +1 or -1.

Given an $m \times n$ matrix A and an m -vector b , we define $P(A, b) = \{x \mid Ax \leq b; x \geq 0\}$.

Baum and Trotter have obtained the following characterisation of t.u. matrices [2].

Theorem 1.1. *Let A be an $m \times n$ matrix of integers; the following statements are equivalent:*

- (1) A is t.u.,
- (2) for each integer m -vector b and for each integer $k \geq 1$, every integer vector x of $P(A, kb)$ is the sum of k integer vectors y^i of $P(A, b)$.

In the next section we will derive a different characterisation which appears

implicitly in Baranyai [1]; for doing this we will mainly use the well-known characterisation of t.u. matrices given by Hoffman and Kruskal [9].

Theorem 1.2. *Let A be an $m \times n$ matrix of integers; the following statements are equivalent.*

- (a) A is t.u.,
- (b) for each integer b , $P(A, b)$ has only integer extreme points.

Another characterisation has been given by Ghouila-Houri [7]. Before stating it, let us define for any real m -vector c a polyhedron $Q(A, c)$ where A is an $m \times n$ matrix of integers by

$$Q(A, c) = \{x \mid [c] \leq Ax \leq [c]; x \geq 0\}.$$

Notice that if c is an integer, then $Q(A, c) = \{x \mid Ax = c; x \geq 0\}$.

Theorem 1.3. [7]. *Let A be an $m \times n$ matrix of integers. Then the following two statements are equivalent.*

- (i) A is t.u.,
- (ii) for each integer m -vector b , every $(0, 1)$ vector x in $Q(A, b)$ is the sum of two $(0, 1)$ vectors y^1, y^2 such that $y^i \in Q(A, \frac{1}{2}b)$ for $i = 1, 2$.

In fact (ii) was given by Ghouila-Houri in the following form: Any subset J of columns may be partitioned into two subsets J_1, J_2 such that for each row s ,

$$\left| \sum_{j \in J_1} a_{sj} - \sum_{j \in J_2} a_{sj} \right| \leq 1.$$

Condition (ii) in Theorem 1.3 is equivalent to the following: For each $(0, 1)$ vector x there exists a $(0, +1, -1)$ vector z such that $z \equiv x \pmod{2}$ and $\alpha_s z = 0$ if $\alpha_s x \equiv 0 \pmod{2}$ or $\alpha_s z = \pm 1$ otherwise ($s = 1, \dots, m$). Here α_s represents row s of A . This condition is quite close to the condition given by Padberg [12] where one considers all $(0, +1, -1)$ vectors x .

A simple proof of Theorem 1.3 based on results of Camion was given by Tamir [13].

Finally Section 3 will contain a discussion of these characterisations in terms of hypergraphs. Relations with balanced and perfect matrices will be mentioned.

2. Properties and characterisation of t.u. matrices

A sequence $(\beta_1, \beta_2, \dots, \beta_k)$ of real nonnegative numbers with $\sum_{i=1}^k \beta_i = 1$ is called bicomposed if by repeatedly grouping any number of equal terms, one may reduce it to a sequence consisting of at most two terms; for instance $(0.3, 0.3, 0.2, 0.1, 0.05, 0.05)$ is bicomposed (it can be reduced to $(0.6, 0.4)$).

We shall need the following preliminary result.

Lemma 2.1. *Let A be an $m \times n$ t.u. matrix, c a real m -vector and k a positive integer. Then any integer x in $Q(A, c)$ is the sum of k integer vectors y^i such that $y^i \in Q(A, (1/k)c)$ for $i = 1, \dots, k$.*

Proof. We will use induction on k ; the result is trivial for $k = 1$, so we assume that it holds for all integers smaller than k . Let x^0 be an integer vector of $Q(A, c)$ and let $b = Ax^0$ (i.e. $\lfloor c \rfloor \leq b \leq \lceil c \rceil$). Then $\bar{y} = (1/k)x^0$ is in $\bar{Q} = Q(A, (1/k)b) \cap \{x \mid x \leq x^0\}$; since (A, I) is t.u. there exists an integer point y^1 in \bar{Q} .

Notice that y^1 is in $Q(A, (1/k)c)$, because $Q(A, (1/k)b) \subseteq Q(A, (1/k)c)$; in order to show this, we have to state the following inequalities:

$$\lfloor (1/k)c \rfloor \leq \lfloor (1/k)b \rfloor, \quad \lfloor (1/k)b \rfloor \leq \lfloor (1/k)c \rfloor. \quad (2.1)$$

If for some s , $b_s = \lfloor c_s \rfloor$, then $\lfloor (1/k)b_s \rfloor = \lfloor (1/k)\lfloor (k/k)c_s \rfloor \rfloor = \lfloor (1/k)c_s \rfloor$ and if $b_s = \lceil c_s \rceil$, then $\lfloor (1/k)b_s \rfloor \geq \lfloor (1/k)c_s \rfloor$. Hence the first inequality (2.1) is established. The second one is proved in the same way.

Now let $y = x^0 - y^1$; clearly $y \geq 0$ and y is integer. Let $b^1 = b - Ay^1$; then $y \in Q(A, b^1)$. By the induction assumption, y is the sum of $k - 1$ integer vectors y^2, \dots, y^k with $y^i \in Q(A, (1/(k-1))b^1)$ for $i = 2, \dots, k$. We have to show that $Q(A, (1/(k-1))b^1) \subseteq Q(A, (1/k)c)$, i.e.,

$$\lfloor (1/(k-1))b^1 \rfloor \leq \lfloor (1/k)c \rfloor, \quad \lfloor (1/(k-1))b^1 \rfloor \geq \lfloor (1/k)c \rfloor. \quad (2.2)$$

Let us establish the first inequality (the second one could be handled in the same way).

$$\begin{aligned} \lfloor (1/(k-1))b^1 \rfloor &= \lfloor (1/(k-1))(b - Ay^1) \rfloor \\ &\leq \lfloor (1/(k-1))(b - \lfloor (1/k)b \rfloor) \rfloor \\ &= \lfloor (1/(k-1))(b + \lceil -(1/k)b \rceil) \rfloor \\ &= \lfloor (1/(k-1))\lceil ((k-1)/k)b \rceil \rfloor \\ &= \lfloor (1/k)b \rfloor \leq \lfloor (1/k)c \rfloor. \end{aligned}$$

The last inequality holds from (2.1). Thus $y^i \in Q(A, (1/k)c)$ for $i = 1, \dots, k$.

Theorem 2.1. *Let A be a t.u. $m \times n$ matrix, b an integer m -vector and $\beta = (\beta_1, \dots, \beta_k)$ a bicomposed sequence. Then any integer vector x in $Q(A, b)$ is the sum of k integer vectors y^i with $y^i \in Q(A, \beta_i b)$ for $i = 1, \dots, k$.*

Remark 2.1. It is not known whether a decomposition of an integer vector x in $Q(A, b)$ into k integer $y^i \in Q(A, \beta_i b)$ exists for any sequence $(\beta_1, \dots, \beta_k)$ of real nonnegative numbers with sum 1.

However given k positive integers n_1, n_2, \dots, n_k with $\sum_{i=1}^k n_i = p$, it follows immediately from Theorem 1.1 that any integer vector x in $P(A, pb)$ is the sum of k integer vectors y^i with $y^i \in P(A, n_i b)$ for $i = 1, \dots, k$.

Proof of Theorem 2.1. Let us assume that β may be reduced to (γ_1, γ_2) and let x be an integer vector of $Q(A, b)$; then $y = \gamma_1 x$ satisfies

$$\lfloor \gamma_1 b \rfloor \leq Ay \leq \lceil \gamma_1 b \rceil, \quad 0 \leq y \leq x.$$

Hence there exists an integer \tilde{y}^1 in $Q(A, \gamma_1 b)$ with $\tilde{y}^1 \leq x$; let $\tilde{y}^2 = x - \tilde{y}^1 \geq 0$. We have $A\tilde{y}^2 = A(x - \tilde{y}^1) \leq b - \lfloor \gamma_1 b \rfloor = \lceil (1 - \gamma_1)b \rceil = \lceil \gamma_2 b \rceil$ and similarly $A\tilde{y}^2 \geq \lfloor \gamma_2 b \rfloor$, i.e., $\tilde{y}^2 \in Q(A, \gamma_2 b)$.

Now starting from (γ_1, γ_2) we will get β by repeatedly dividing each term of the sequence into a given number of equal terms. Let us examine the general step of this procedure.

More precisely suppose we start from a sequence $\delta = (\delta_1, \dots, \delta_p)$ and we get a sequence $\epsilon = (\epsilon_1, \dots, \epsilon_p)$ by splitting δ_1 into r equal terms $\epsilon_1, \dots, \epsilon_r$ (i.e. $\epsilon_i = \delta_1/r$ for $i = 1, \dots, r$).

If we assume that we have obtained a decomposition of x into p integer vectors z^1, \dots, z^p with

$$z^i \in Q(A, \delta_i b) \quad \text{for } i = 1, \dots, p$$

we have to show that z^1 can be decomposed into r integer vectors w^1, \dots, w^r with

$$w^i \in Q(A, \epsilon_i b) \quad \text{for } i = 1, \dots, r.$$

According to Lemma 2.1, z^1 is the sum of r integer vectors w^i such that $w^i \in Q(A, (\delta_1/r)b) = Q(A, \epsilon_i b)$ for $i = 1, \dots, k$. So we will finally get the required decomposition of x .

Theorem 2.2. Let A be an $m \times n$ matrix of integers; the following two statements are equivalent.

- (i) A is t.u.,
- (ii) for each integer m -vector b and for each integer $k \geq 1$, every integer x in $Q(A, b)$ is the sum of k integer vectors y^i such that $y^i \in Q(A, (1/k)b)$ for $i = 1, \dots, k$.

Proof. From Theorem 2.1, (i) implies (ii) because $(1/k, \dots, 1/k)$ is bicomposed. The converse is a direct consequence of Theorem 1.3.

3. t.u., balanced and perfect matrices

As can be expected, similar characterisations can be given for balanced and for perfect matrices. We will use the terminology of Berge [4] for hypergraphs.

A matrix of zeros and ones is *balanced* if it does not contain any odd submatrix with row sums and column sums equal to 2 [5].

It was shown by Berge [5] that A is balanced iff $P(A', e')$ has only integer extreme points (for each submatrix A' of A and for each compatibly dimen-

sioned vector e' with all components equal to 1). A hypergraph associated with a balanced matrix will be called *balanced* as well.

Perfect matrices can be defined as cliques-nodes matrices of perfect graphs [11]; it has been shown that a matrix is perfect iff $P(A, e)$ has only integer extreme points where e has all components equal to 1.

The results of the previous sections on t.u. matrices can be summarized as follows.

Theorem 3.1 (total unimodularity). *Let A be a $m \times n$ matrix of integers; the following statements are equivalent.*

(tu0) A is t.u.,

(tu1) for each integer m -vector b and for each integer $k \geq 1$, every integer vector x of $P(A, kb)$ is the sum of k integer vectors of $P(A, b)$,

(tu2) for each integer m -vector b and for each integer $k \geq 1$, every integer vector x of $Q(A, b)$ is the sum of k integer vectors y^i satisfying

$$c \leq Ay^i \leq d, \quad 0 \leq y^i$$

where $c_s = \lfloor b_s/k \rfloor$ and $d_s = \lceil b_s/k \rceil \leq 1$ ($s = 1, \dots, m$).

For balanced matrices, we have:

Theorem 3.2 (balanced matrices). *Let A be a $m \times n$ matrix of zeros and ones; the following conditions are equivalent.*

(b0) A is balanced,

(b1) for each integer $k \geq 1$ and for each submatrix A' , of A , every integer vector x of $P(A', ke')$ (where e' is a compatibly dimensioned vector with all components equal to 1) is the sum of k integer vectors of $P(A', e')$,

(b2) for each integer $k \geq 1$ and for each integer m -vector b , every integer vector x of $Q(A, b)$ is the sum of k integer vectors y^i satisfying

$$c \leq Ay^i \leq d, \quad 0 \leq y^i$$

where

$$\begin{array}{lll} c_s = \lfloor b_s/k \rfloor & \text{and} & d_s = \lceil b_s/k \rceil \quad \text{if } b_s \leq k, \\ c_s = 1 & \text{and} & d_s = \infty \quad \text{if } b_s > k. \end{array}$$

The corresponding statements for perfect matrices are:

Theorem 3.3 (perfect matrices). *Let A be an $m \times n$ matrix of zeros and ones; the following conditions are equivalent.*

(p0) A is perfect.

(p1) for each integer $k \geq 1$, every integer vector x of $P(A, ke)$ (where e has components equal to 1) is the sum of k integer vectors of $P(A, e)$.

(p2) for each integer m -vector b and for each $k \geq \max_s b_s$, each integer vector x

of $Q(A, b)$ is the sum of k integer vectors y^i satisfying

$$c \leq Ay^i \leq d, \quad 0 \leq y^i$$

where $c_s = \lfloor b_s/k \rfloor$, $d_s = \lceil b_s/k \rceil \leq 1$ ($s = 1, \dots, m$).

If A is a $(0, 1)$ matrix, we may associate with it a hypergraph H defined as follows: each row of A corresponds to an edge of H and each column of A to a node of H ; $a_{sj} = 1 \Leftrightarrow$ edge E_s contains node j . If A is t.u., hypergraph H will be called a *unimodular* hypergraph. Condition (tu2) has an immediate interpretation in terms of node coloring in H . We define an *equitable* k -coloring (see [4, 14]) of the nodes of H to be a partition of the node set X of H into k subsets X_1, \dots, X_k such that for each edge E_s and for each c ($1 \leq c \leq k$)

$$\lfloor |E_s|/k \rfloor \leq |X_c \cap E_s| \leq \lceil |E_s|/k \rceil.$$

Now if in (ii) we choose an integer $x = (1, 1, \dots, 1)$ (this implies that $b_s = \sum_j a_{sj} = |E_s|$), then any decomposition of x into k integer $y^i \in Q(A, (1/k)b)$ will define an equitable k -coloring of the nodes of H (y^i will be the characteristic vector of X_i , i.e. $y^i_j = 1$ if node j is in X_i). So we have:

Corollary 3.1 [14]. *A unimodular hypergraph has an equitable k -coloring of nodes for each $k \geq 2$.*

Furthermore Theorem 1.3 applied to hypergraphs means that: A hypergraph H is unimodular iff each subhypergraph H' has an equitable bicoloring. (H' is obtained from H by deleting some nodes, i.e. by removing some columns of A .)

Remark 3.1. By replacing in Theorem 3.1 condition (tu1) by a condition (tu1)' obtained by setting $k = 2$ and by considering only $(0, 1)$ vectors as in Theorem 1.3, we would get the same characterisation of unimodular hypergraphs:

(tu2)' for each integer m -vector b , every $(0, 1)$ vector x in $P(A, 2b)$ is the sum of two $(0, 1)$ vectors y^1, y^2 with $y^i \in P(A, b)$.

In terms of unimodular hypergraphs, this condition can be expressed as follows: Given any integer m -vector b , any subset S of nodes verifying $|S \cap E_s| \leq 2b_s$ for each edge E_s may be partitioned into two subsets S_1, S_2 such that $|S_i \cap E_s| \leq b_s$ for each edge E_s and for $i = 1, 2$. If S is fixed, choosing the smallest possible values for b_s , i.e. $b_s = \lceil |S \cap E_s|/2 \rceil$, (S_1, S_2) defines an equitable bicoloring of the subhypergraph H' of H constructed on S .

One should observe however that if $k > 2$, a k -coloring (X_1, \dots, X_k) satisfying $|X_c \cap E_s| \leq \lceil |E_s|/k \rceil$ for each edge E_s and for each c ($1 \leq c \leq k$) is not necessarily equitable.

Proof of Theorem 3.2. Condition (b2) for a balanced hypergraph H means that

for each integer $k \geq 2$, it has a *good* k -coloring, i.e., a partition of the node set into k subsets X_1, X_2, \dots, X_k such that for each edge E_s and for each color i ($1 \leq i \leq k$) $| \{i \mid X_i \cap E_s \neq \emptyset\} | = \min(k, |E_s|)$ (Berge [6] has shown that $(b0) \Rightarrow (b2)$).

Condition $(b2)'$ obtained by considering only $(0, 1)$ vectors x and by setting $k = 2$ would mean that any subhypergraph of H has a good bicoloring. It is also known that $(b2) \Rightarrow (b0)$.

Similarly condition $(b1)'$ means that in every partial hypergraph each 2-independent set of nodes is the sum of two 1-independent sets of nodes (a p -independent set of nodes is an assignment of a value $p(x) \in \{0, 1, \dots, p\}$ to each node x in such a way that for each edge E_s $\sum_{x \in E_s} p(x) \leq p$; if $p = 1$, we have a usual independent (or stable) set).

It is known that multiplying the nodes of a balanced hypergraph still gives a balanced hypergraph; since a balanced hypergraph with rank k (i.e. $\max_s |E_s| = k$) has strong chromatic number k [4], we have $(b0) \Rightarrow (b1)$.

For showing $(b1) \Rightarrow (b0)$, we may assume that A is not balanced. Hence there exists a submatrix A' of A such that $P(A', e')$ has a noninteger extreme point x^* . Since all entries in A' are integers, we may choose an integer $k \geq 1$ such that $x' = kx^*$ is an integer point in $P(A', ke')$. By assumption x' is the sum of k integers y^i with $y^i \in P(A', (1/k)ke') = P(A', e')$; so $x^* = (1/k) \sum_{i=1}^k y^i$. Since x^* is an extreme point, we must have $x^* = y^1 = \dots = y^k$; this is not possible because x^* is not integer while all y^i 's are. Hence A is balanced.

Proof of Theorem 3.3. As for balanced matrices, the equivalence of $(p0)$ and $(p2)$ is easily established: $(p2)$ restricted to $(0, 1)$ -vectors x and to $k = \max_s b_s$ means that for any subhypergraph H' of H , the strong chromatic number is equal to the rank in H' ; or if A is the clique-matrix of a perfect graph G , then for any subgraph G' , the chromatic number $\gamma(G')$ is equal to $\omega(G')$ the maximum cardinality of a clique in G' .

For showing $(p1) \Rightarrow (p0)$, we may proceed exactly as in the proof of Theorem 3.2.

Conversely, if A is the clique-matrix of a perfect graph, then any integer vector x of $P(A, ke)$ can be considered as the incidence vector of a subset of nodes in a graph kG obtained from G by multiplying its nodes by k . kG and its subgraphs are perfect. Since $x \in P(A, ke)$, no clique of the corresponding subgraph of kG has more than k nodes, hence it can be colored with at most k colors; this gives the required decomposition of x .

Another way of showing that $(p0) \Rightarrow (p1)$ would be to apply a recent result of Baum and Trotter [3] concerning the decomposition property of points in lower comprehensive polyhedra. Essentially if A is the clique-node incidence matrix of a perfect graph, then for any nonnegative integer vector w ,

$$\min\{1y \mid yA \geq w; y \geq 0 \text{ integer}\} = \min\{1y \mid yA \geq w; y \geq 0\}.$$

This is equivalent to saying that for any positive integer k any integer x in $P(A, ke)$ is the sum of k integer y^i with $y^i \in P(A, e)$.

As discussed by Lovasz [10], hypergraphs associated with perfect matrices have duals which are normal. Condition (p1) restricted to $k=2$ and to $(0, 1)$ vectors means that in H itself (but not necessarily in its partial hypergraphs) each 2-independent set of nodes is the sum of two 1-independent sets of nodes. This is equivalent to the condition stated by Lovasz ([10], lemma 1): each 2-matching of H^* (dual of H) is the sum of two 1-matchings.

References

- [1] Z. Baranyai, "The edge-coloring of complete hypergraphs, I", *Journal of Combinatorial Theory* B26 (1979) 276–294.
- [2] S. Baum and L.E. Trotter, "Integer rounding and polyhedral decomposition for totally unimodular systems", in: R. Henn, B. Korte and W. Oetli, eds., *Arbeitstagung über Operations Research und Optimierung* (Springer, Berlin, 1978) pp. 15–23.
- [3] S. Baum and L.E. Trotter, "Integer rounding for polymatroid and branching optimization problems, Bonn Univ. O.R./Econ. Inst. Techn. Report no. 78120-OR.
- [4] C. Berge, *Graphs and hypergraphs* (North-Holland, Amsterdam, 1973).
- [5] C. Berge, "Balanced matrices", *Mathematical Programming* 2 (1972) 19–31.
- [6] C. Berge, "Notes sur les bonnes colorations d'un hypergraphe", *Cahiers du Centre d'Etudes de Recherche Opérationnelle* 15 (1973) 219–223.
- [7] A. Ghouila-Houri, "Caractérisation des matrices totalement unimodulaires", *Comptes rendus de l'Académie des Sciences, Paris* 254 (1962) 1192.
- [8] M. Gondran, "Matrices totalement unimodulaires", *Bulletin EDF, Série C*, 1 (1973) 55–74.
- [9] A.J. Hoffman and J.B. Kruskal, "Integral boundary points of convex polyhedra", in: H.W. Kuhn and A.W. Tucker, eds., *Linear inequalities and related systems* (Princeton Univ. Press, Princeton, NJ, 1956) pp. 223–246.
- [10] L. Lovasz, "Minimax theorems for hypergraphs", in: C. Berge and D. Ray-Chaudhuri, eds., *Hypergraph Seminar* (Springer, Berlin, 1974) pp. 111–126.
- [11] M.W. Padberg, "Characterisations of totally unimodular, balanced and perfect matrices", in: B. Roy, ed., *Combinatorial programming: Methods and applications* (D. Reidel Publ., Dordrecht, Holland, 1975) pp. 275–284.
- [12] M.W. Padberg, "A note on the total unimodularity of matrices", *Discrete Mathematics* 14 (1976) 273–278.
- [13] A. Tamir, "On totally unimodular matrices", *Networks* 6 (1976) 373–382.
- [14] D. de Werra, "Equitable colorations of graphs", *Revue Française d'Informatique et de Recherche Opérationnelle* R-3 (1971) 3–8.