1. Easy Interger LPs

Consider the linear integer program

$$\begin{array}{lll} \mathcal{IP} & \text{minimize} & c^T x \\ & \text{subject to} & Ax = b,\ 0 \le x\ ,\ \text{and } x \in \mathbb{Z}^n\ , \end{array}$$

where $c \in \mathbb{R}^m$, $A \in \mathbb{Z}^{m\times n}$, and $b \in \mathbb{Z}^m$. The LP relaxation of $\mathcal{IP}$ is the linear program

$$\begin{array}{lll} \mathcal{IP}_{\mathsf{relax}} & \text{minimize} & c^T x \\ & \text{subject to} & Ax = b,\ 0 \le x\ . \end{array}$$

We say that the a solution to $\mathcal{IP}_{\mathsf{relax}}$ is integral if all of its components are integer, that is, it is feasible for $\mathcal{IP}$. Note that if the solution to $\mathcal{IP}_{\mathsf{relax}}$ is integral, then it must also be the solution to $\mathcal{IP}$. In this section we ask the question

"When is the solution to the LP relaxation $\mathcal{IP}_{\mathsf{relax}}$ integral?"

Our answer requires us to more closely examine the nature of solutions to $\mathcal{IP}_{\mathsf{relax}}$ and how they are represented. Recall from the Fundamental Theorem of Linear Programming that if an LP has a solution, then it must have a basic feasible solution, or equivalently a vertex solution. Let us review what this means.

Definition 1.1. *A point x in the closed convex set $C \subset \mathbb{R}^n$ is said to be an extreme point of C if whenever $u, v \in C$ are such that $x \in [u, v]$, then either $x = u$ or $x = v$, where*

$$[u, v] = \{(1-\lambda)u + \lambda v \ |0 \le \lambda \le 1\}$$

is the line segment connecting u and v.

If C is polyhedral convex, then we call the extreme points the vertices of C.

Theorem 1.1. *Consider the the polyhedron $\Omega = \{x\ |Ax = b,\ 0 \le x\}$, where $m < n$, $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$. We will assume that A is surjective (*$\operatorname{rank}(A) = m$*). Then $x \in \mathbb{R}^n$ is a vertex of Ω if and only if $x \in \Omega$ and there is an index set $\mathcal{B} \subset \{1, \ldots, n\}$ with $|\mathcal{B}| = m$, $\{i\ |x_i > 0\} \subset \mathcal{B}$ and $A_{\mathcal{B}}$ is nonsingular, where $A_{\mathcal{B}}$ is the matrix formed from the columns of A in the index set $\mathcal{B}$.*

Proof. Let us first assume that x is a vertex of Ω. Set $\mathcal{B} = \{i\ |x_i > 0\}$ and $\mathcal{I} = \{i\ |x_i = 0\}$. Assume there is a $z_{\mathcal{B}} \in \operatorname{Nul}(A_{\mathcal{B}})\setminus\{0\}$. Since $x_{\mathcal{B}} > 0$, there is an $\epsilon > 0$ such that $x_{\mathcal{B}} \pm t z_B > 0$ whenever $-\epsilon \le t \le \epsilon$. Let $\bar{x}, \bar{z} \in \mathbb{R}^n$ be such that

$$\bar{x}_i = \begin{cases} (x_{\mathcal{B}})_i & i \in \mathcal{B}, \\ 0 & i \notin \mathcal{B} \end{cases} \quad \text{and } \bar{z}_i = \begin{cases} (z_{\mathcal{B}})_i & i \in \mathcal{B} \\ 0 & i \notin \mathcal{B}\ . \end{cases}$$

Then $A(\bar{x} \pm \epsilon\bar{z}) = b$ and $0 \le \bar{x} \pm \epsilon\bar{z}$, that is, $\bar{x} \pm \epsilon\bar{z} \in \Omega$ and $x = \frac{1}{2}(\bar{x} - \epsilon\bar{z}) + \frac{1}{2}(\bar{x} + \epsilon\bar{z})$ contradicting the fact that x is a vertex of Ω. Hence $A_{\mathcal{B}}$ cannot have a null-space. In particular, we have $|\mathcal{B}| \le m$. If $|\mathcal{B}| = m$ we are done, so we assume that $|\mathcal{B}| < m$. Since $\operatorname{Nul}(A_{\mathcal{B}}) = \{0\}$, the columns of $A_{\mathcal{B}}$ are linearly independent. Since $\operatorname{rank}(A) = m$, one can find an index set of $m - |\mathcal{B}|$ other columns of A say $\mathcal{C}$ such that $A_{\hat{\mathcal{B}}}$ is nonsingular, where $\hat{\mathcal{B}} = \mathcal{B} \cup \mathcal{C}$. Reset $\mathcal{B}$ to $\hat{\mathcal{B}}$ to obtain the result.

Next we assume that $x \in \Omega$ and there is an index $\mathcal{B} \subset \{1, \ldots, n\}$ with $|\mathcal{B}| = m$, $\{i\ |x_i > 0\} \subset \mathcal{B}$ and $A_{\mathcal{B}}$ is nonsingular. We need to show that x is a vertex. If x is not a vertex, then there exist $x_1, x_2 \in \Omega$ with $x_1 \neq x \neq x_2$ and $x = \frac{1}{2}(x_1 + x_2)$, or equivalently, $x_2 = x + \frac{1}{2}z$ and $x_1 = x - \frac{1}{2}z$ where $z = x_2 - x_1 \neq 0$. In particular, $x \pm \frac{1}{2}z \in \Omega$. Let $j \in \mathcal{I} = \{i \,|x_i = 0\}$. Since $(x \pm \frac{1}{2}z)_j \geq 0$, we must have $z_j = 0$ for all $j \in \mathcal{I}$. Since $z \neq 0$ and $z_{\mathcal{I}} = 0$, we have $z_{\mathcal{B}} \neq 0$. Moreover,

$$b = A(x + \frac{1}{2}z) = A_{\mathcal{B}}(x_{\mathcal{B}} + \frac{1}{2}z_{\mathcal{B}}) = b + \frac{1}{2}A_{\mathcal{B}}z_{\mathcal{B}},$$

so that $z_{\mathcal{B}} \in \mathrm{Nul}\,(A_{\mathcal{B}})$. This contradicts the assumption that $A_{\mathcal{B}}$ is nonsingular. Therefore x is a vertex of Ω. □

Recall that the index set $\mathcal{B}$ is call the basis, and the variables $x_j,\ j \in \mathcal{B}$ are called the basic variables. Using the notation introduced in Theorem 1.1, for any basis $\mathcal{B} \subset \{1, 2, \ldots, n\}$ we can write

$$\begin{aligned} b &= A_{\mathcal{B}}x_{\mathcal{B}} + A_{\mathcal{N}}x_{\mathcal{N}} \\ z &= c_{\mathcal{B}}^T x_{\mathcal{B}} + c_{\mathcal{N}}x_{\mathcal{N}}\ , \end{aligned}$$

where $\mathcal{N} = \{1, 2, \ldots, n\} \setminus \mathcal{B}$ is the index set of the non-basic variables. Since $A_{\mathcal{B}}$ is nonsingular, we can multiply through by $A_{\mathcal{B}}^{-1}$ and rearrange to get

$$x_{\mathcal{B}} = A_{\mathcal{B}}^{-1}b - A_{\mathcal{B}}^{-1}A_{\mathcal{N}}x_{\mathcal{N}} \tag{1}$$

$$z = \mathcal{B}^T A_{\mathcal{B}}^{-1}b + (c_{\mathcal{N}} - A_{\mathcal{N}}A_{\mathcal{B}}^{-1}c_{\mathcal{B}})^T x_{\mathcal{N}}\ . \tag{2}$$

This is called the *dictionary* associated with the basis $\mathcal{B}$. Setting $x_{\mathcal{N}} = 0$ gives $x_{\mathcal{B}} = A_{\mathcal{B}}^{-1}b$ which is called the basic solution associated with this basis. If, in addition, we have $0 \leq A_{\mathcal{B}}^{-1}b$, then this is called a *basic feasible solution.* Hence, by Theorem 1.1, the basic feasible solutions correspond precisely to the vertices of the polyhedron Ω. Finally, the augmented matrix associated with the system (1)-(2) is called the simplex tableau associated with the basis $\mathcal{B}$.

With this algebraic structure in mind, we wish to determine conditions under which the optimal basic feasible solution has all integer components. That is, if $\mathcal{B}$ is the optimal basis, under what conditions is the vector

$$A_{\mathcal{B}}^{-1}b$$

a vector of integers. We begin by first studying the question of when the solution of a square system in integral. That is, given $A \in \mathbb{Z}^{n\times n}$ nonsingular and $b \in \mathbb{Z}^n$, under what conditions is the solution to the system $Ax = b$ integral? We approach this question using a very classical method known as Cramer's Rule which requires the use of determinants.

The Leibniz formula for the determinant of an $n \times n$ matrix A is

$$\det(A) = \sum_{\sigma \in \mathcal{S}_n} \mathrm{sgn}\,(\sigma) \prod_{i=1}^{n} A_{i\sigma(i)}, \tag{3}$$

where $\mathcal{S}_n$ is the set of all permutations of the integers $\{1, 2, ..., n\}$. These permutations are functions that reorder this set of integers. The element in position i after the reordering σ is denoted $\sigma(i)$. For example, for n = 3, the original sequence $[1, 2, 3]$ might be reordered to $[2, 3, 1]$, with $\sigma(1) = 2,\ \sigma(2) = 3,\ \sigma(3) = 1$. The set of all such permutations (also known as the symmetric group on n elements) is denoted $\mathcal{S}_n$. For each permutation σ, $\mathrm{sgn}\,(\sigma)$ denotes the signature of σ; it is $+1$ for even σ and -1 for odd σ. Evenness or oddness can be defined as follows: the permutation is even (odd) if the new sequence can be obtained by an even number (odd, respectively) of switches of numbers. For example, starting from $[1, 2, 3]$ and switching the positions of 2 and 3 yields $[1, 3, 2]$, switching once more yields $[3, 1, 2]$, and finally, after a total of three (an odd number) switches, $[3, 2, 1]$ results. Therefore $[3, 2, 1]$ is an odd permutation. Similarly, the permutation $[2, 3, 1]$ is even since

$$[1, 2, 3] \to [2, 1, 3] \to [2, 3, 1],$$

an even number of switches.

The identity permutation is the unique element $\iota \in \mathcal{S}_n$ for which $\iota(i) = i$ for all $i = 1, 2, \ldots, n$. Note that there are zero switches. We say that ι is even so that its signature is 1, $\operatorname{sgn}(\iota) = 1$. Observe that if I is the identity matrix, then

$$\det(I) = \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} I_{i\,\sigma i} = \operatorname{sgn}(\iota) = 1.$$

Given σ_1, $\sigma_2 \in \mathcal{S}_n$, we can use composition $\sigma = \sigma_1 \circ \sigma_1$ to obtain another element of $\mathcal{S}_n$. The discussion above on the signature of a permutation tells us that $\operatorname{sgn}(\sigma_1 \circ \sigma) = \operatorname{sgn}(\sigma_1)\operatorname{sgn}(\sigma_2)$. Therefore, if $\tilde{A}$ is the matrix obtained from A by permuting its columns using the permutation $\pi \in \mathcal{S}_n$, then Leibniz's formula (3) tells us that

$$\begin{aligned}
\det(\tilde{A}) &= \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} A_{i\sigma(\pi(i))} \\
&= \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma)[\operatorname{sgn}(\pi)]^2 \prod_{i=1}^{n} A_{i\sigma(\pi(i))} \\
&= \operatorname{sgn}(\pi) \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma)\operatorname{sgn}(\pi) \prod_{i=1}^{n} A_{i\sigma(\pi(i))} \\
&= \operatorname{sgn}(\pi) \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma \circ \pi) \prod_{i=1}^{n} A_{i(\sigma\circ\pi)(i)} \\
&= \operatorname{sgn}(\pi) \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} A_{i\sigma(i)} \\
&= \operatorname{sgn}(\pi) \det(A).
\end{aligned}$$

Since we can also write $\tilde{A} = AP_\pi$, where $P_\pi \in \mathbb{R}^{n\times n}$ is the permutation matrix corresponding to π, this gives us the formula

$$\det(AP_\pi) = \operatorname{sgn}(\pi) \det(A).$$

Taking $A = I$ and using the fact that $\det(I) = 1$, we obtain

$$\det(P_\pi) = \operatorname{sgn}(\pi) \det(I) = \operatorname{sgn}(\pi) \quad \forall\, \pi \in \mathcal{S}_n. \tag{4}$$

Therefore, for every permutation matrix $P \in \mathbb{R}^{n\times n}$ and matrix $A \in \mathbb{R}^{n\times n}$, we have

$$\det(AP) = \det(P) \det(A). \tag{5}$$

If $n = 2$, then $\mathcal{S}_2$ only has 2 elements:

$$[1, 2] \xrightarrow{\iota} [1, 2] \text{ and } [1, 2] \xrightarrow{\sigma} [2, 1],$$

where $\text{sgn}(\iota) = 1$ and $\text{sgn}(\sigma) = -1$. Therefore, given a 2×2 matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ its determinent is

$$\det\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = ad - bc \ .$$

We have the following basic result on how to compute the determinant recursively.

Theorem 1.2 (Laplace's Formula for the Determinant)**.** *Suppose $A \in \mathbb{R}^{n \times n}$. For every pair $i, j \in \{1, 2, \ldots, n\}$, define the matrix $A[i, j]$ to be the $(n-1) \times (n-1)$ submatrix of A obtain by deleting the ith row and the jth column. The for each $i_0, j_0 \in \{1, 2, \ldots, n\}$,*

$$\det(A) = \sum_{i=1}^{n} a_{ij_0}(-1)^{(i+j_0)} \det(A[i, j_0]) = \sum_{j=1}^{n} a_{i_0 j}(-1)^{(i_0+j)} \det(A[i_0, j]) \ .$$

The terms $C_{ij} = (-1)^{(i+j)} \det(A[i, j])$ are called the cofactors of the matrix A and the transpose of the matrix whose ijth component is C_{ij} is called the *classical adjoint* of A denoted $\text{adj}(A) = [C_{ij}]^T$. The determinant satisfies the following properties.

Theorem 1.3 (Properties of the Determinant)**.** *Let $A, B \in \mathbb{R}^{n \times n}$.*

(1) $\det(A) = \det(A^T)$.

(2) *The determinant is a multi-linear function of its columns (rows). That is, if $A = [A_{\cdot 1}, A_{\cdot 2}, \ldots, A_{\cdot n}]$, where $A_{\cdot j}$ is the jth column of A ($j = 1, \ldots, n$), then for any vector $b \in \mathbb{R}^n$ and scalar $\lambda \in \mathbb{R}$*

$$\det([A_{\cdot 1}, \ldots, A_{\cdot j} + \lambda b, \ldots, A_{\cdot n}]) = \det([A_{\cdot 1}, \ldots, A_{\cdot j}, \ldots, A_{\cdot n}]) + \lambda \det([A_{\cdot 1}, \ldots, b, \ldots, A_{\cdot n}]) \ .$$

(3) *If any two columns (rows) of A coincide, then* $\det(A) = 0$.

(4) *For every $j_1, j_2 \in \{1, \ldots, n\}$ with $j_1 \neq j_2$ and $\lambda \in \mathbb{R}$,*

$$\det(A) = \det([A_{\cdot 1}, \ldots, A_{\cdot j_1} + \lambda A_{\cdot j_2}, \ldots, A_{\cdot n}]) \, .$$

(5) *If A is singular, then* $\det(A) = 0$.

Proof. (1) This follows immediately from Laplace's formula for the determinant in Theorem 1.2.

(2) This follows immediately from Laplace's formula:

$$\begin{aligned}
\det([A_{\cdot 1}, \ldots, A_{\cdot j} + \lambda b, \ldots, A_{\cdot n}]) &= \sum_{i=1}^{n} (a_{ij} + \lambda b_i)(-1)^{(i+j)} \det(A[i, j]) \\
&= \sum_{i=1}^{n} a_{ij}(-1)^{(i+j_0)} \det(A[i, j]) + \lambda \sum_{i=1}^{n} b_i(-1)^{(i+j_0)} \det(A[i, j]) \\
&= \det([A_{\cdot 1}, \ldots, A_{\cdot j}, \ldots, A_{\cdot n}]) + \lambda \det([A_{\cdot 1}, \ldots, b, \ldots, A_{\cdot n}]) \ .
\end{aligned}$$

(3) The permutation $\pi_{ij} \in \mathcal{S}_n$ that interchanges i and j ($\pi_{ij}(i) = j,\ \pi_{ij}(j) = i,\ \pi_{ij}(k) = k \, \forall \, k \neq i, j$) is odd since $\text{sgn}(\pi_{ij}) = 2\,|i - j| - 1$ as long as $i \neq j$. Therefore, by (4), the

permutation matrix P_{ij} which interchanges the columns $i \neq j$ has $\det(P_{ij}) = -1$. Now suppose that column i equals column $j \neq i$ in the matrix $A \in \mathbb{R}^{n\times n}$, then, by (5), $\det(A) = \det(AP_{ij}) = \det(A)\det(P_{ij}) = -\det(A)$. Hence $\det(A) = 0$.
(4) This follows immediately from Parts (2) and (3).
(5) If A is singular, then its columns are linearly dependent. That is, there is a non-trivial linear combination of its columns that give zero, or equivalently, there is some column j_0 that is a linear combination of the remaining columns, $A_{\cdot j_0} = \sum_{j\neq j_0} \lambda_j A_{\cdot j}$. Therefore, by Parts (2) and (3),

$$\begin{aligned} \det(A) &= \det\left([A_{\cdot 1}, \ldots, A_{\cdot (j_0-1)}, \sum_{j\neq j_0} \lambda_j A_{\cdot j}, A_{\cdot (j_0+1)}, \ldots, A_{\cdot n}]\right) \\ &= \sum_{j\neq j_0} \lambda_j \det\left([A_{\cdot 1}, \ldots, A_{\cdot (j_0-1)}, A_{\cdot j}, A_{\cdot (j_0+1)}, , \ldots, A_{\cdot n}]\right) \\ &= 0 \ . \end{aligned}$$

□

We will also need two further properties fo the determinant. These appear in the next theorem whose proof is omitted.

Theorem 1.4. *Let $A, B \in \mathbb{R}^{n\times n}$, $C \in \mathbb{R}^{m\times n}$, and $D \in \mathbb{R}^{m\times m}$. Then the following two formulas hold:*

$$\det(AB) = \det(A)\det(B) \tag{6}$$

$$\det\left(\begin{bmatrix} A & 0 \\ C & D \end{bmatrix}\right) = \det(A)\det(D) \ . \tag{7}$$

Note that (5) is a special case of (6). As an application of (6) we compute the determinant of A^{-1} when it exists:

$$1 = \det(I) = \det(AA^{-1}) = \det(A)\det(A^{-1}),$$

whenever A is nonsingular. That is, $\det(A^{-1}) = \det(A)^{-1}$.

We now return to the our discussion of the system $Ax = b$ and Cramer's Rule. Cramer's Rule states that if A is nonsingular, then the unique solution to the system $Ax = b$ is given componentwise by

$$x_j = \frac{\det(A_j(b))}{\det(A)}, \ \ j = 1, 2, \ldots, n,$$

where the matrix $A_j(b)$ is obtained from A by replacing the jth column of A by the vector b. The proof of Cramer's Rule follows easily from the properties of the determinant. Indeed, if $\bar{x}$ is the unique solution to the system $Ax = b$, then $b = A\bar{x} = \sum_{j=1}^{n} \bar{x}_j A_{\cdot j}$. Therefore, by

Parts (2) and (3) of Theorem 1.3,

$$\begin{aligned}
\det(A_j(b)) &= \det\left([A_{\cdot 1},\dots,A_{\cdot(j-1)},b,A_{\cdot(j+1)},\dots,A_{\cdot n}]\right)\\
&= \det\left([A_{\cdot 1},\dots,A_{\cdot(j-1)},\sum_{i=1}^{n}\bar{x}_iA_{\cdot i},A_{\cdot(j+1)},\dots,A_{\cdot n}]\right)\\
&= \sum_{i=1}^{n}\bar{x}_i\det\left([A_{\cdot 1},\dots,A_{\cdot(j-1)},A_{\cdot i},A_{\cdot(j+1)},\dots,A_{\cdot n}]\right)\\
&= \bar{x}_i\det\left([A_{\cdot 1},\dots,A_{\cdot(i-1)},A_{\cdot i},A_{\cdot(i+1)},\dots,A_{\cdot n}]\right)\\
&= \bar{x}_i\det(A)
\end{aligned}$$

giving Cramer's Rule.

Let us examine the expressions $\det(A_j(b))$ using Laplace's formula for the determinant:

$$\det(A_j(b)) = \sum_{i=1}^{n} b_iC_{ij} = C_{\cdot j}^T b,$$

where $C_{\cdot j}$ is the jth row of the classical adjoint $\operatorname{adj}(A)$. That is,

$$\bar{x} = \frac{1}{\det(A)}\operatorname{adj}(A)b.$$

Since this expression is valid for all choices of $b \in \mathbb{R}^n$, we must have

$$A^{-1} = \frac{1}{\det(A)}\operatorname{adj}(A).$$

We call this the adjoint representaton of the inverse.

Next suppose that the matrix A has only integer components. Then Leibniz's formula (3) tells us that $\det(A)$ is also integer and that $\operatorname{adj}(A)$ is integer since each cofactor of A, $C_{ij} = (-1)^{(i+j)}\det(A[i,j])$ is integer. Hence, from the adjoint representation of the inverse, we see that A^{-1} must have all integer components if $\det(A) \in \{-1,1\}$. This motivates the following definition and theorem.

Definition 1.2. *A matrix $A \in \mathbb{R}^{n\times n}$ is said to be unimodular if* $\det(A) \in \{-1,1\}$.

Theorem 1.5. *Let $A \in \mathbb{Z}^{n\times n}$ be non-singular. Then the following are equivalent:*

(1) *The solution to $Ax = b$ is integral for every $b \in \mathbb{Z}^n$.*
(2) $A^{-1} \in \mathbb{Z}^{n\times n}$.
(3) *A is unimodular.*
(4) *A^{-1} is unimodular.*

Proof. If every solution to $Ax = b$ are integral for every $b \in \mathbb{Z}^n$, then, in particular, the solutions to $Ax = e_i$ is integral for each unit coordinate vector e_i, $i = 1,2,\dots,n$. But these are just the columns of A^{-1}. Hence $A^{-1} \in \mathbb{Z}^{n\times n}$.

If $A^{-1} \in \mathbb{Z}^{n\times n}$, then $1 = det(I) = \det(AA^{-1}) = \det(A)\det(A^{-1})$. But both $\det(A)$ and $\det(A^{-1})$ are integral from (3) since both are elements of $\mathbb{Z}^{n\times n}$. Therefore, both $\det(A)$ and $\det(A^{-1})$ can only take the values ± 1. Thus, in particular, both A and A^{-1} are unimodular. $\square$

This result indicates that unimodularity is the perfect tool for analyzing the integrality of solutions to integral systems of equations. The first step is to appropriately extend the notion of unimodularity to the context of the integer linear programs $\mathcal{IP}$ introduced at the beginning of this section. Here the underlying matrix $A \in \mathbb{Z}^{m \times n}$ is nolonger square. Nonetheless, recall that the solutions to $\mathcal{IP}$ obtained from the simplex algorithm are necessarily basic feasible solutions having the form described in (1) with $x^*_{\mathcal{B}} = A_{\mathcal{B}}^{-1} b \geq 0$. That is, the optimal solution x^* comes from solving the square system $A_{\mathcal{B}} x_{\mathcal{B}} = b$ for $x^*_{\mathcal{B}}$ using the optimal basis $\mathcal{B}$ and then setting the non-basic entries to zero, $x^*_{\mathcal{N}} = 0$. Hence, to guarentee that every vector generated in this way is integral, we should assume that $A_{\mathcal{B}}$ is unimodular for every possible choice of basis $\mathcal{B}$. This leads to the following somewhat stronger idea.

Definition 1.3. *A matrix $A \in \mathbb{Z}^{m \times n}$ is said to be totally unimodular if every nonsingular square submatrix of A is unimodular.*

Note that an immediate consequence of this definition is that if A is totally unimodular, then the entries in A can only take values from the set $\{0, \pm 1\}$.

Theorem 1.6. *Let $A \in \mathbb{Z}^{m \times n}$ with A surjective, i.e.* $\mathrm{Ran}\,(A) = \mathbb{R}^m$. *For each $b \in \mathbb{Z}^m$, define the convex polyhedron $\Omega(b) = \{x \in \mathbb{R}^n \mid Ax = b,\ 0 \leq x\}$. If A is totally unimodular, then, for every $b \in \mathbb{Z}^m$, all of the vertices of $\Omega(b)$ are integral.*

Proof. Let $b \in \mathbb{Z}^m$. Then, for every basis $\mathcal{B}$, $A_{\mathcal{B}}^{-1} b$ is integral by Theorem 1.5. In particular, every basic feasible solution, or equivalently, every vertex of $\Omega(b)$ is integral. Since b was chosen arbitrarily from $\mathbb{Z}^m$, we have that every vertex of $\Omega(b)$ is integral for every $b \in \mathbb{Z}^m$. Note that the proof does not require that $\Omega(b) \neq \emptyset$, since in this case the result holds trivially. □

This is a fine result that seems to answer the question we posed at the beginning of this section. There is only one problem. Are there any totally unimodular matrices of interest in practice? Just how common and useful are these beasts? Would we recognize one if it passed us on the street? As a partial answer to these questions, we provide the following two results.

Theorem 1.7. *Let $A \in \mathbb{R}^{m \times n}$. Then the following statements are equivalent.*

(1) *A is totally unimodular.*
(2) *A^T is totally unimodular.*
(3) *The matrix $[A\ I]$ is totally unimodular.*

Proof. The equivalence follows immediately from the fact that the determinant of a matrix equals the determinant of its transpose. It remains to establish the equivalence of (a) and (b).

Clearly, if $[A\ I]$ is totally unmodular, then A must be totally unimodular since every nonsingular square submatrix of A is a submatrix of $[A\ I]$. Therefore, we need only prove the reverse implication. Let B be a nonsingular square submatrix of $[A\ I]$. Then there is a permutation of the rows (P_1) and columns (P_2) of B so that it has the form

$$P_1 B P_2 = \begin{bmatrix} B_{11} & 0 \\ B_{21} & I \end{bmatrix},$$

where the B_{11} and B_{21} contain distinct entries from A. Therefore, $\det(B) = \pm\det(B_{11})$. Since B is nonsingular, so is B_{11}. Hence B_{11} is a nonsingular square submatrix of A, it is unimodular. Therefore, $\det(B) = \pm 1$. ☐

Theorem 1.8. *Let $A \in \mathbb{Z}^{m\times n}$ and let a_{ij} denote the ijth entry of A. If A satisfies the following three conditions, then A is totally unimodular.*

(1) *$a_{ij} \in \{0, \pm 1\}$ for all ij.*
(2) *Every column of A has at most two non-zero entries.*
(3) *The rows of A can be partitioned into two index sets I_1 and I_2 such that*
   (a) *If a column has two entries of different signs, then the indices of the rows corresponding to these non-zero entries must be in the same index set.*
   (b) *If a column has two entries of the same sign, then the indices of the rows corresponding to these non-zero entries must be in different index sets.*

Proof. Suppose that the result is false, that is, there is a matrix A satisfying these three conditions that is not totally unimodular. Let B be any nonsingular submatrix of A of smallest degree that is not unimodular, and let k be this degree. Clearly, $k > 1$ since if $k = 1$, then $B = \pm 1$ contradicting our standing hypotheses. Clearly, B can have no zero column, and due to the smallest degree requirement, B can have no column with a single non-zero entry (why?). Hence every column of B has precisely two non-zero entries. Define the vector $v \in \mathbb{Z}^k$ by

$$v_i = \begin{cases} 1 & i \in I_1 \\ -1 & i \in I_2 \end{cases} .$$

Then

$$(v^T B)_j = \sum_{i\in I_1} a_{ij} - \sum_{i\in I_2} a_{ij} = 0,\ j = 1, 2, \ldots, k ,$$

that is $v^T B = 0$ which contradicts the assumption that B is nonsingular. Hence no such B can exist so that A is totally unimodular. ☐

We will soon see that there are a number of important classes of problems for which the matrix A satisfies he conditions given in Theorem 1.8. We can also use Theorem 1.7 to extend Theorem 1.6 beyond integer LPs of the form $\mathcal{IP}$.

Theorem 1.9. *If $A \in \mathbb{Z}^{m\times n}$ is totally unimodular, then, for every $b \in \mathbb{Z}^m$, every vertex of the polyhedron $\{x \in \mathbb{R}^n \mid Ax \le b,\ 0 \le x\}$ is integral. In particular, this implies that for every $b \in \mathbb{Z}^m$ and $c \in \mathbb{R}^n$, the LP*

$$\begin{array}{ll} \textit{minimize} & c^T x \\ \textit{subject to} & Ax \le b,\ 0 \le x \end{array}$$

has an integral solution whenever a solution exists.

Proof. Since A is totally unimodular, Theorem 1.7 tells us that the matrix $[A\ I]$ is totally unimodular. It was shown in Math 407 (The Fundamental Representation Theorem for Vertices) that x is a vertex of $\{x \in \mathbb{R}^n \mid Ax \le b,\ 0 \le x\}$ if and only if there is a vector of slacks s such that $\binom{x}{s}$ is a vertex of the set $\left\{\binom{x}{s} \,\middle|\, [A\ I]\binom{x}{s} \le b,\ 0 \le \binom{x}{s}\right\}$. Since Theorem 1.6

tells us that all of the vertices of this latter set are integral, the vertices of the former are integral as well.

The final statement follows from the Fundamental Theorem of Linear Programming which states that if an LP has a solution, then it must have a vertex solution. □