Integer Lattices

In this section we consider rational matrices and vectors, and will be interested in the consistency of and finding (characterizing) some (all) integral solutions $x \in \mathbb{Z}^n$ of a system of linear equations

$$Ax = b \tag{1}$$

for a given matrix $A \in \mathbb{Q}^{m \times n}$ and vector $b \in \mathbb{Q}^m$. Let us denote by a^j, $j = 1, .., n$ the column vectors of A, i.e.,

$$A = [a^1, a^2, ..., a^n].$$

The *lattice* $\langle A \rangle$ generated by the columns of A is defined as the set of vectors obtainable as integral combinations of the columns of A, i.e.,

$$\langle A \rangle = \{Ax \mid x \in \mathbb{Z}^n\}.$$

Thus, the basic problem of consistency of (1) over $\mathbb{Z}^n$ can equivalently be restated as checking if $b \in \langle A \rangle$, or not.

To be able to describe a solution to the above problem, we need to consider certain matrix operations. First of all, let us assume in the sequel that A is of full row rank (since otherwise we can find the dependent rows and eliminate those from the above system of equations, in polynomial time, e.g., by Gaussian elimination).

The following three basic operations on the columns of a matrix A are called *elementary column operations*:

(i) exchange columns a^i and a^j, for some $i \neq j$;

(ii) replace column a^i by $-a^i$;

(iii) replace column a^i by $a^i + \lambda a^j$ for some $j \neq i$ and $\lambda \in \mathbb{Z}$.

Let us call an integral matrix $U \in \mathbb{Z}^{n \times n}$ *unimodular* if $\det U = \pm 1$, and let us observe a few easy facts. Clearly, U is a unimodular matrix if and only if both U and U^{-1} are integral:

$$U \cdot U^{-1} \; = \; I \quad \Rightarrow \quad \det(U) \cdot \det(U^{-1}) \; = \; 1.$$

If both $\det(U)$ and $\det(U^{-1})$ are integers, then ...

Furthermore, we have the following properties:

Lemma 1 *If A' is obtained from A by a series of elementary column operations, then*

(a) $A' = AU$, *where U is a unimodular matrix;*

(b) A *can also be obtained from* A' *by a series of elementary column operations, namely, by applying the inverse operations in reverse order, i.e.,* $A = A'U^{-1}$*;*

(c) $\langle A' \rangle = \langle A \rangle$.

Proof.

$$\left[\mathbf{a}^1, \ldots, -\mathbf{a}^i, \ldots, \mathbf{a}^n\right] \quad = \quad \left[\mathbf{a}^1, \ldots, \mathbf{a}^i, \ldots, \mathbf{a}^n\right] \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & -\mathbf{1} & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix}$$

$$\left[\mathbf{a}^1, \ldots, \mathbf{a}^j, \ldots, \mathbf{a}^i, \ldots, \mathbf{a}^n\right] \quad = \quad \left[\mathbf{a}^1, \ldots, \mathbf{a}^i, \ldots, \mathbf{a}^j, \ldots, \mathbf{a}^n\right] \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & & \mathbf{1} & & \\ & & & \ddots & & & \\ & & \mathbf{1} & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix}$$

$$\left[\mathbf{a}^1, \ldots, \mathbf{a}^i + \lambda \cdot \mathbf{a}^j, \ldots, \mathbf{a}^j, \ldots, \mathbf{a}^n\right] \quad = \quad \left[\mathbf{a}^1, \ldots, \mathbf{a}^i, \ldots, \mathbf{a}^j, \ldots, \mathbf{a}^n\right] \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & \mathbf{1} & & & & \\ & & & \ddots & & & \\ & & \lambda & & \mathbf{1} & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix}$$

□

Let us call a matrix $A' = [B, 0] \in \mathbb{Q}^{m\times n}$, where $B = (b_{ij})$ is a square submatrix, the *Hermite normal form* of A (remember, A is assumed to be of full row rank), if A' can be obtained from A by a series of elementary column operations, and if the following properties hold for B:

(I) $0 \leq b_{ij} < b_{ii}$ for all i and $j < i$;

(II) $b_{ij} = 0$ for all i and $j > i$.

Theorem 1 *Every rational matrix can be brought to Hermite normal form by a series of elementary column operations.*

Proof. It follows by the definitions that if $[B, 0]$ is a Hermite normal form of A, and $\lambda \in \mathbb{Z}_+$, then $[\lambda B, 0]$ is a Hermite normal form of λA. Thus, we can assume that A integral matrix (and is of full row rank).

Let us apply elementary column operations in stages, where in stage k we assume that the first $k-1$ rows are already satisfying conditions (I) and (II). Clearly, there are at most m stages.

Let D denote the matrix we obtained after the first $k-1$ stages. We shall work with the last $n-k+1$ columns of D at the beginning of stage k, and therefore these operations will not change the fact that

$$d_{ij} = 0 \text{ for all } i \leq k-1 \text{ and } j \geq k.$$

Let $(d^0_{kj} \mid j = 1, ..., n)$ denote the elements in the k-th row of D, and let us first multiply the columns by -1 if needed (operation (ii)), so that we have $d^0_{kj} \geq 0$ for all $j \geq k$, and set $t = 0$ as the iteration counter.

Let us then apply the following steps, iteratively:

- permute the columns (by applying operation (i)) so that

  $$d^t_{kk} \geq d^t_{kk+1} \geq \cdots \geq d^t_{kn} \geq 0.$$

  Let j^t be the index of the last positive entry, i.e., $d^t_{kj^t}$ is the smallest nonzero entry in row k. If $j^t = k$, then STOP; otherwise

- set for all $k \leq j < j^t$ (operation (iii))

  $$d^t_{kj} = d^t_{kj} - \left\lfloor \frac{d^t_{kj}}{d^t_{kj^t}} \right\rfloor d^t_{kj^t},$$

  and set $t = t+1$.

We can see that $d^t_{kj} \geq 0$ for all $j \geq k$ and all t, and that upon termination row k has the form $(*, ..., *, d_{kk}, 0, \ldots, 0)$ for some $d_{kk} > 0$ value ($d_{kk} = 0$ would imply that row k of A is not linearly independent from the other rows, contradicting our full row rank assumption).

Let us then apply operation (iii) for all $j = 1, ..., k-1$:

$$d_{kj} = d_{kj} - \left\lfloor \frac{d_{kj}}{d_{kk}} \right\rfloor d_{kk}$$

after which row k will satisfy both conditions (I) and (II).

To finish the proof we need only to show that the above procedure terminates in a finite number of steps. To see this assume that d^t_{kk} denotes the largest element among d_{kj}, $j = k, k+1, ..., n$ as right after the first permutation step, before the other steps. We claim that

$$d^{t+2}_{kk} \leq \frac{1}{2} d^t_{kk}$$

holds for all iteration t (for which iteration $t+1$ is still has to be carried out). This claim follows by the observations that

$$d^{t+2}_{kk} = d^{t+1}_{kj^{t+1}} \leq d^{t+1}_{kk} = d^t_{kj^t},$$

and that

$$d^{t+1}_{kj^{t+1}} \leq d^t_{kk} - d^t_{kj^t}.$$

Since we assumed that A is an integral matrix, all these numbers are integers, and thus the number of iterations is limited by $2 \log_2 d^0_{kk}$, and hence the procedure is indeed finite. □

Since adding a zero vector to a vector set does not change the generated lattice, the following easy fact follows:

Lemma 2 *If $[B, 0]$ is the Hermite normal form of A, then*

(d) $\langle A\rangle = \langle B\rangle$.

Proof.

$$\begin{aligned}
\langle A\rangle = &\ \{Ax \mid x \in \mathbb{Z}^n\} \\
= &\ \left\{AUU^{-1}x \mid x \in \mathbb{Z}^n\right\} \\
= &\ \left\{[B, 0] \cdot U^{-1}x \mid x \in \mathbb{Z}^n\right\} \\
= &\ \left\{[B, 0]\begin{pmatrix} y \\ z \end{pmatrix} \mid \begin{pmatrix} y \\ z \end{pmatrix} \in \mathbb{Z}^{m+(n-m)}\right\} \\
= &\ \{By \mid y \in \mathbb{Z}^m\} \\
= &\ \langle B\rangle
\end{aligned}$$

□

Let us show next that the Hermite normal form of a rational matrix is unique (hence justifying the use of "*normal form*" in its name). The uniqueness will follow form the following, somewhat stronger statement.

Theorem 2 *Let $A \in \mathbb{Z}^{m\times n}$ and $A' \in \mathbb{Z}^{m\times n'}$ be full row rank integral (rational) matrices, and let $[B, 0]$ and $[B', 0]$ be their respective Hermite normal forms. Then, $\langle A\rangle = \langle A'\rangle$ if and only if $B = B'$.*

Proof. If $B = B'$, then we have $\langle A\rangle = \langle B\rangle = \langle B'\rangle = \langle A'\rangle$ by easy fact (d).

Let us thus assume that $\langle A\rangle = \langle A'\rangle$ and assume indirectly that $B \neq B'$, where $B = (b_{ij})$ and $B' = (b'_{ij})$. Let us then choose the smallest row index i for which there exists a column index j such that $b_{ij} \neq b'_{ij}$. We can assume without any loss of generality that $b_{ij} \geq b'_{ij}$, and let $\Delta = (b_{kj} - b'_{kj} \mid k = 1, ..., m)$ be the difference between b^j and b'^j, the j-th columns of B and B'.

Since $\Lambda = \langle B\rangle = \langle A\rangle = \langle A'\rangle = \langle B'\rangle$ by our assumptions and easy fact (d), Δ is the difference of two vectors of the lattice Λ, and hence itself belongs to Λ. Since $\Lambda = \langle B\rangle$, we have

$$\Delta = \sum_{k=1}^{m} \beta_k b^k$$

for some integers $\beta_k \in \mathbb{Z}$, $k = 1, ..., m$. Because we have $\Delta_k = 0$ for $k < i$ by our choice of index i, we must have $\beta_k = 0$ for $k < i$[1]. Furthermore, we have $b_{ik} = 0$ for $k > i$ by (II), thus it follows that $\Delta_i = b_{ij} - b'_{ij} = \beta_i b_{ii}$. On the other hand, we also have $0 < b_{ij} - b'_{ij} < b_{ii} - b'_{ij} \leq b_{ii}$ by (II) and by our assumptions and choice of index i. The resulting inequalities $0 < \beta_i b_{ii} < b_{ii}$ however contradict the integrality of β_i, proving that our indirect assumption was incorrect, i.e. that $B = B'$. □

Corollary 1 *Every rational matrix has a unique Hermite normal form.* □

Lemma 3 *If $M \in \mathbb{Z}^{m\times m}$ is a nonsingular matrix and $\mu = \det M$, then we have $\langle M\rangle \supseteq \langle \mu I\rangle$.*

Proof. We have $\mu M^{-1} \in \mathbb{Z}^{m\times m}$, and $\mu I = M\,[\mu M^{-1}]$, i.e., the columns of μM^{-1} provide the integral combinations of the columns of M yielding the columns of μI. □

Corollary 2 *A matrix $U \in \mathbb{Z}^{m\times m}$ is unimodular if and only if $\langle U\rangle = \langle I\rangle$.*

Proof. Assume first that $\det U = 1$. Then, since for every integral matrix $A \in \mathbb{Z}^{m\times n}$ we have

$$\langle A\rangle \subseteq \langle I\rangle = \mathbb{Z}^m,$$

which together with Lemma 3 imply that $\langle I\rangle \supseteq \langle U\rangle \supseteq \langle(\det U)I\rangle = \langle I\rangle$, because we have $\det U = 1$ assumed.

For the reverse direction, assume that $\langle U\rangle = \langle I\rangle$, i.e., that $I = UM$ for some integral matrix $M \in \mathbb{Z}^{m\times m}$, from which

$$1 = \det I = (\det U)(\det M)$$

follows. Since both U and M are integral, their determinants are integers, too. Thus, the above equality implies $\det U = \pm 1$, i.e., that U is indeed unimodular. □

[1] WHY? Think of induction for $k = 1, ..., i-1$ and the facts that $b_{kk} > 0$ by (I)!

The following theorem show that the Hermite normal form can always be efficiently computed (see e.g., Kannan and Bachem 1979; Domich 1983; Domich, Kannan and Trotter 1985).

Theorem 3 *The Hermite normal form of a rational matrix can be computed in polynomial time.*

Proof. We can assume[2] that the input is a full row rank, integral matrix $A \in \mathbb{Z}^{m\times n}$. Let M be an $m \times m$ non-singular submatrix of A (since A is of full row rank, such an M exists), and let $\mu = \det M$. Then we have $\langle A\rangle \supseteq \langle M\rangle$ by definition, and hence

$$\langle [A \mid \mu I]\rangle \;=\; \langle A\rangle.$$

Then by Theorem 2 we can compute the Hermite normal form of A by computing that of $A' = [A \mid \mu I]$. Let us do that the same way as in the proof of Theorem 1, with the change that in each main iteration we use the last m columns and operation (iii) to keep all entries in A' with an absolute value not larger than μ. Then all computations in the course of the algorithm will involve integers representable on $O(\log \mu)$ bits, and the procedure will terminate in $O(m \log \mu)$ iterations. Since μ is bound by a polynomial in the input size of A, the theorem follows. □

Corollary 3 *If $A \in \mathbb{Q}^{m\times n}$ is a rational matrix and $[B, 0]$ is its Hermite normal form, then $[B, 0] = AU$ for some unimodular matrix $U \in \mathbb{Z}^{n\times n}$, where the size of U is polynomially limited in the size of A.* □

Corollary 4 *For $A \in \mathbb{Q}^{m\times n}$ and $b \in \mathbb{Q}^m$ the problem of deciding the consistency of the system of equations $Ax = b$ over $x \in \mathbb{Z}^n$ has a good characterization.*

[2] WHY?

Proof. Compute the Hermite normal form $AU = [B, 0]$, which we can do in polynomial time and which is of polynomial size in the size of the input, according to Corollary 3. Now, if $B^{-1}b \in \mathbb{Z}^m$, then $x = U(B^{-1}b, 0) \in \mathbb{Z}^n$ is a solution. Otherwise, $B^{-1}b$ has a fractional component, proving that $b \notin \langle A \rangle$. □

Corollary 5 *Given a rational system of equations $Ax = b$, we can find in polynomial time integer vectors $y^j \in \mathbb{Z}^n$, $j = 0, 1, ..., t$, where $t = n - rank(A)$, such that*

$$\{x \mid Ax = b,\ x \in \mathbb{Z}^n\} \ = \ \{y^0 + \lambda_1 y^1 + \cdots + \lambda_t y^t \mid \lambda_1, \ldots, \lambda_t \in \mathbb{Z}\}.$$

Proof. Let $AU = [B, 0]$ be the Hermite normal form of A, which according to Corollary 3 can be computed in polynomial time. Then, we have

$$\left[y^0, y^1, \ldots, y^t\right] \ = \ U \left[\begin{array}{c|c} B^{-1}b & 0 \\ \hline 0 & I \end{array}\right].$$

□

Theorem 4 (Mordell, 1969) *A rational system of equations $Ax = b$ has an integral solution if and only if $Ax \equiv b \pmod{M}$ is consistent for all $M \in \mathbb{Z}$.* □

Theorem 5 (Kronecker, 1884) *Either the rational system $Ax = b$ of equations has an integral solution, or there exists a rational vector $y \in \mathbb{Q}^m$ such that $y^T A \in \mathbb{Z}^n$ and $y^T b \notin \mathbb{Z}$, but not both.*

Proof. If $Ax = b$ has an integral solution $x \in \mathbb{Z}^n$, and $y \in \mathbb{Q}^m$ is such that $y^A \in \mathbb{Z}^n$, then we have $y^T b = y^T(Ax) = (y^T A)x \in \mathbb{Z}$, showing that both conditions cannot hold simultaneously.

For the other direction, assume that $Ax = b$ has no integral solution, and let us compute the Hermite normal form $AU = [B, 0]$ of A. In this case, as in the proof of Corollary 4, we must have a fractional component in $B^{-1}b$. Let y be the row of B^{-1}, corresponding to this fractional component. Then we have $y^T b \notin \mathbb{Z}$, while $y^T A = y^T A(UU^{-1}) = (y^T AU)U^{-1} = (y^T[B, 0])U^{-1} \in \mathbb{Z}^n$, since $y^T B$ is a unit vector, and U^{-1} is an integral matrix. □