Integer Points in Polyhedra Gennady Shmonin

Lattices and Hermite normal form

February 17, 2009

1 Lattices

Let $B = \{b_1, b_2, \ldots, b_k\}$ be a set of linearly independent vectors in n-dimensional Euclidean space $\mathbb{R}^n$. The set of the form

$$\Lambda(B) := \Big\{ \sum_{i=1}^{k} \lambda_i b_i : \lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{Z} \Big\}$$

is called a *lattice* with *basis* B (or *generated by* B) and k is the *dimension* of $\Lambda(B)$. If $k = n$, we say that the lattice is *full-dimensional*. We shall mostly concentrate on full-dimensional lattices; otherwise we may apply an appropriate linear transformation and identify $\mathrm{span}(B)$ with $\mathbb{R}^k$.

Yet, the basis can compactly be represented as a square $n \times k$-matrix (we also denote it by B) with vectors $b_1, b_2, \ldots, b_n$ as columns. Then we can write

$$\Lambda(B) = \{Bx : x \in \mathbb{Z}^k\}. \tag{1}$$

We shall often interchange between different notation without mentioning this explicitly, but it should not cause any confusion; thus, B is either a set of vectors $b_1, b_2, \ldots, b_k$, or a matrix with columns $b_1, b_2, \ldots, b_k$, depending on the context. Trivially, the vectors $b_1, b_2, \ldots, b_k$ are linearly independent if and only if the matrix B has full column rank, i.e., $\mathrm{rank}(B) = k$, and they span $\mathbb{R}^n$ if and only if B has full row rank.

It is natural to ask about sets of the form (1) when vectors in B are not necessarily linearly independent. Is it still a lattice? In other words, are there linearly independent vectors generating the same set $\Lambda(B)$? In general, the answer is 'no': just consider the one-dimensional vectors $b_1 = [1]$ and $b_2 = [\sqrt{2}]$ and the corresponding set $\Lambda(b_1, b_2) = \{\lambda_1 + \lambda_2\sqrt{2} : \lambda_1, \lambda_2 \in \mathbb{Z}\}$. However, if we restrict ourselves onto rational vectors only, then the answer is 'yes', as we shall see a little later. We shall also see that lattices are exactly the discrete subgroups of $\mathbb{R}^n$ (a group Λ is called *discrete* if there is a neighbourhood of the origin containing no elements from $\Lambda \setminus \{0\}$—clearly, this property does not hold for our example $\{\lambda_1 + \lambda_2\sqrt{2} : \lambda_1, \lambda_2 \in \mathbb{Z}\}$).

The set of integral vectors $\mathbb{Z}^n$ is a full-dimensional lattice in $\mathbb{R}^n$, generated by the unit vectors $e_1, e_2, \ldots, e_n$. But the same lattice can also be generated by the vectors $e_1 + e_2, e_2, e_3, \ldots, e_n$ (since the sum $\lambda_1 e_1 + \lambda_2 e_2$, where $\lambda_1, \lambda_2 \in \mathbb{Z}$, can be rewritten as $\lambda_1(e_1 + e_2) + (\lambda_2 - \lambda_1)e_2$, and the coefficients remain integral; for the converse, we rewrite $\lambda_1(e_1 + e_2) + \lambda_2 e_2$ as $\lambda_1 e_1 + (\lambda_1 + \lambda_2)e_2$). Thus, the basis of a lattice is not unique. However, all the bases of a given lattice are equivalent modulo "unimodular transformations". Recall that an integral square matrix U is called *unimodular* if $|\det(U)| = 1$. The following simple facts will be needed.

Lemma 1. *Let U be a unimodular matrix. Then*
(a) the inverse U^{-1} is also unimodular;
(b) x is an integral vector if and only if Ux is an integral vector.

Proof. Since U is an integral matrix, all cofactors of U are integers. It follows that all entries of U^{-1} are also integers, as $|\det(U)| = 1$. Finally, $UU^{-1} = I$ implies $\det(U)\det(U^{-1}) = 1$, whence $|\det(U^{-1})| = 1$. This proves part (a). It is obvious that if x is an integral vector, then Ux is also an integral vector. The converse follows now from part (a), since $x = U^{-1}(Ux)$. □

Now, we can establish the connexion between different bases of the same lattice. For the sake of both completeness and consistency, in the following lemma we also consider general-form matrices of full row rank, where do not claim that the set $\Lambda(B)$ is a lattice—$\Lambda(B)$ is just the group generated by the columns of B, i.e., the set defined as in (1), where the vectors are not necessarily linearly independent.

Lemma 2. *Let B and B' be matrices. If $B = B'U$ for some unimodular matrix U, then $\Lambda(B) = \Lambda(B')$. Moreover, if B and B' have full column rank and $\Lambda(B) = \Lambda(B')$, then $B = B'U$ for some unimodular matrix U.*

Proof. Let U be a unimodular matrix and suppose that $B = B'U$. Then for every vector $y = Bx$ with $x \in \mathbb{R}^n$, we have $y = B'(Ux)$, and x is integral if and only if Ux is integral. This implies $\Lambda(B) = \Lambda(B')$.

If B and B' have full column rank and $\Lambda(B) = \Lambda(B')$, then $B \subseteq \Lambda(B')$, and therefore, $B = B'V$ for some integral matrix V. Similarly, $B' \subseteq \Lambda(B)$, which implies that $B' = BW$ for some integral matrix W. It follows that $B = BWV$, and since B has full column rank, $WV = I$. Particularly, $V = W^{-1}$ and $\det(V)\det(W) = 1$. But V and W are integral matrices, and the last equality is only possible when $|\det(W)| = |\det(V)| = 1$. □

2 Hermite normal form

The following *elementary column operations* are particular unimodular transformations of a matrix $B = [b_1, b_2, \ldots, b_n]$:

(1) swap two columns in B: $b_i \longleftrightarrow b_j$ $(i \neq j)$;

(2) multiply a column by -1: $b_i \longleftarrow (-b_i)$;

(3) add an integer multiple of a column to another column: $b_i \longleftarrow b_i + \alpha b_j$ $(i \neq j, \alpha \in \mathbb{Z})$.

We explicitly specify appropriate unimodular matrices for each of the elementary column operations. Thus, swapping columns b_i and b_j in matrix B is equivalent to multiplying B by the unimodular matrix $U = [u_{kl}]$, which is obtained from the identity matrix by swapping its i-th and j-th columns. Precisely, $u_{kk} = 1$ for all $k \neq i, j$, $u_{ij} = u_{ji} = 1$, and $u_{kl} = 0$ otherwise. Multiplying a column b_i of matrix B by -1 is equivalent to multiplying B with the unimodular matrix $U = [u_{kl}]$, which is obtained from the identity matrix by multiplying the i-th column by -1; hence, $u_{kk} = 1$ for all $k \neq i$, $u_{ii} = -1$, and $u_{kl} = 0$ otherwise. Lastly, adding α times column b_j to column b_i in matrix B is equivalent to multiplying B with the unimodular matrix U, which is obtained from the identity matrix by α times j-th column to the i-th column, i.e., $u_{kk} = 1$ for all k, $u_{ji} = \alpha$, and $u_{ij} = 0$ otherwise.

We say that a matrix B of full row rank is in *Hermite normal form* if it has the form $B = [H \mid 0]$, where $H = [h_{ij}]$ is a square matrix such that

(1) $h_{ij} = 0$ for $i < j$ (i.e., H is lower-triangular);

(2) $0 \leqslant h_{ij} < h_{ii}$ for $i > j$ (i.e., H is non-negative and each row has a unique maximum entry, which is on the main diagonal)

Particularly, matrix H is non-singular. Now, we show that any matrix can be brought into Hermite normal form by applying an appropriate sequence of elementary column operations. Since the elementary column operations are actually unimodular transformations of a matrix, the group generated by the columns of the matrix is invariant under these operations; in other words, if we had transformed the matrix into a matrix in Hermite normal form, we also proved that this group can be generated by linearly independent vectors, and therefore, is a lattice.

Theorem 3 (Existence of Hermite normal form)**.** *Each rational matrix of full row rank can be brought into Hermite normal form by a sequence of elementary column operations.*

Proof. Let B be a rational matrix of full row rank. Without loss of generality, we may assume that B is integral; otherwise, $B = \frac{1}{\delta}B'$, where δ is the least common multiple of all denominators in B, and we proceed with matrix B'. We describe an algorithm converting B into a matrix in Hermite normal form. This algorithm constructs a sequence of matrices $B_1, B_2, \ldots$, where

$$B_k = \begin{bmatrix} H_k & 0 \\ C_k & D_k \end{bmatrix},$$

where H_k is a $k \times k$-matrix in Hermite normal form, and the matrix B_{k+1} is obtained from the matrix B_k as follows.

Let $d_1, d_2, \ldots, d_{n-k}$ be the entries in the first row of D_k. By permuting some columns and multiplying some columns by -1, we may assure that they all are non-negative. Moreover, there is at least one non-zero entry, since B has full row rank. If $d_i > d_j$ are two non-zero entries in the first row of D_k, we add $-\left\lfloor \frac{d_i}{d_j} \right\rfloor$ times the j-th column to the i-th column of D_k. All the entries in the first row of D_k remain non-negative but their total sum strictly decreases. (In fact, we execute the Euclidean algorithm to compute the greatest common divisor of d_i and d_j, and therefore, terminate if both d_i and d_j are integers.) Therefore, by repeating this procedure we end up with exactly one non-zero entry, say d, in the first row of D_k, which after swapping the columns is located in the first column. It remains to ensure that all entries in the first row of C_k are non-negative and smaller than d. To do so, we add $-\left\lfloor \frac{c_i}{d} \right\rfloor$ times the $(k+1)$-th column to the i-th column in B_k, for each $i = 1, 2, \ldots, k-1$ (clearly, this does not affect the entries of H_k). □

Due to unimodularity of elementary column operations, we can derive the following corollary.

Corollary 3a. *Let B be a matrix of full row rank. Then there is a unimodular matrix U such that the matrix BU is in Hermite normal form.*

Corollary 3a implies that every rational lattice has a basis in Hermite normal form. Moreover, if B is a rational matrix of full row rank, then the group generated by B, $\Lambda(B)$, is a lattice. In the next section we state these facts in a slightly more general form.

In fact, the proof of Theorem 3 yields an algorithm to compute Hermite normal form of a matrix. But is this algorithm polynomial? Unfortunately, not yet. The problem is that the entries may grow exponentially during the execution of the algorithm (recall the Gaussian elmination method with cross-multiplication). Later, we shall consider this question once again and modify the algorithm, so that it runs in polynomial time.

3 Sublattices

As we mentioned before, Theorem 3 shows, as a side-effect, that any group generated by rational vectors is a lattice, i.e., it can be generated by linearly independent vectors. In this section, we state this result in a slightly more general form.

Let $\Lambda = \Lambda(B)$ be a full-dimensional lattice in $\mathbb{R}^n$. It follows directly from Lemma 2 that the value $|\det(B)| > 0$ is independent of the particular choice of a basis. Thus, we may define the *determinant* of lattice Λ as $\det(\Lambda) := |\det(B)|$.

Let $\Lambda' = \Lambda(B')$ be another full-dimensional lattice. If $\Lambda' \subseteq \Lambda$, we say that Λ' is a *sublattice* of Λ. In this case, $B' = BV$ for some integral matrix V. The value

$$D(\Lambda, \Lambda') := |\det(V)| = \frac{|\det(B')|}{|\det(B)|} = \frac{\det(\Lambda')}{\det(\Lambda)}$$

is then a positive integer and is called the *index* of Λ' in Λ. In particular, if B' is an integral matrix, then $|\det(\Lambda')|$, the determinant of lattice Λ', is the index of Λ' in $\mathbb{Z}^n$.

Lemma 4. *Let Λ' be a sublattice of a lattice Λ (both Λ and Λ' are full-dimensional). Then*

$$D\Lambda \subseteq \Lambda' \subseteq \Lambda,$$

where D is the index of Λ' in Λ and $D\Lambda'$ denotes the scaled lattice $\Lambda(DB') = \{Dx : x \in \Lambda\}$.

Proof. The second inclusion being trivial, we prove that $D\Lambda \subseteq \Lambda$. Let B and B' be bases of Λ and Λ', respectively. It is clear that DB is a basis of $D\Lambda$. Since $\Lambda' \subseteq \Lambda$, we have $B' = BV$ for some integral matrix V and $D = |\det(V)| > 0$. But then $B = B'V^{-1}$, DV^{-1} is an integral matrix, and therefore, $DB \subseteq \Lambda'$. This implies $D\Lambda \subseteq \Lambda'$. □

Theorem 5 (Bases in Hermite normal form)**.** *Let Λ' be a sublattice of Λ.*

(a) For every basis B of Λ, there is a unique basis B' of Λ' such that $B' = BH$, where H is a matrix in Hermite normal form.

(b) For every basis B' of Λ', there is a unique basis B of Λ such that $B = B'H'$, where H' is a matrix in Hermite normal form.

Proof. Let B be a basis of lattice Λ and choose any basis B' of lattice Λ'. Since $\Lambda' \subseteq \Lambda$, we have $B' = BV$ for some integral matrix V. By Corollary 3a, there is a unimodular matrix U such that $H = VU$ is in Hermite normal form, whence $B'U = BVU = BH$. By Lemma 2, $B'U$ is also a basis of Λ' and Part (a) is proved.

For Part (b), let B' be a basis of lattice Λ' and choose an arbitrary basis B of lattice Λ. By Lemma 4, we have $D\Lambda \subseteq \Lambda'$, where D is the index of B' in B, and therefore $B = \frac{1}{D}B'V$ for some integral matrix V. Again, there is a unimodular matrix U such that $H = \frac{1}{D}VU$ is in Hermite normal form. Hence, $BU = \frac{1}{D}B'VU = B'H$ is the required basis of B. □

A particular case of Theorem 5 is that of *rational lattices*, i.e., lattices generated by rational vectors: they always have a basis in Hermite normal form.

Theorem 6. *Let Λ be a full-dimensional lattice with basis $B \in \mathbb{R}^{n\times n}$ and let $B' = \left[b'_1, b'_2, \ldots, b'_k\right]$ be a matrix composed of some vectors $b'_1, b'_2, \ldots, b'_k \in \Lambda$ that span $\mathbb{R}^n$. Then the set group generated by B', $\Lambda(B')$, is a lattice.*

Proof. Again, $B' = BV$ for some integral matrix $V \in \mathbb{Z}^{n\times k}$. By Corollary 3a, there is a unimodular matrix U such that $H = VU$ is in Hermite normal form. By Lemma 2, $\Lambda(B') = \Lambda(B'U)$ and $B' = BVU = BH$. But H is in Hermite normal form, and therefore, has only n non-zero columns. The same is therefore true for $B'U$, implying that $\Lambda(B'U)$ is a lattice. □

For rational bases, this means that $\Lambda(B)$ is a lattice whenever B is a rational matrix.