The Dual Lattice, Integer Linear Systems and Hermite Normal Form

1 Dual Lattice

In the first Lecture, we saw that lattices can be viewed equivalently in two different ways, i.e. as discrete additive subgroup of $\mathbb{R}^n$ or as an additive subgroup of $\mathbb{R}^n$ with linearly independent generators. In this section, we show a final equivalent viewpoint which relates the discreteness of a lattice to the existence of an appropriate dual.

Definition 1 (Dual Lattice) *For an additive subgroup $G \subseteq \mathbb{R}^n$, we define the dual lattice of G to be*

$$G^* = \{\mathbf{x} \in \text{span}(G) : \langle \mathbf{x}, \mathbf{y} \rangle \in \mathbb{Z}, \mathbf{y} \in G\}.$$

The dual lattice is a very useful object for proving geometric inequalities about the original lattice. In particular, we can often use dual vectors to provide an effeciently checkable proof of some property of the original lattice. In the following lemma, we show that dual vectors can provide a simple witness for proving lower bounds on the length of the shortest non-zero vector.

Lemma 2 *Let $G \subseteq \mathbb{R}^n$ be an additive subgroup of rank $k \geq 1$. Then if there exists linearly independent vectors $\mathbf{b}_1^*, \ldots, \mathbf{b}_k^* \in G^*$, we have that*

$$\lambda_1(G) \geq \min_{1 \leq i \leq k} \frac{1}{\|\mathbf{b}_i^*\|_2}$$

In particular, if $\dim(G^*) = \dim(G)$, then G is a lattice.

Proof: Take $\mathbf{x} \in G \setminus \{\mathbf{0}\}$. Since $\mathbf{x} \in \text{span}(\mathbf{b}_1^*, \ldots, \mathbf{b}_k^*)$ and $\mathbf{x} \neq \mathbf{0}$, there exists $j \in [k]$ such that $\langle \mathbf{x}, \mathbf{b}_j^* \rangle \neq 0$. Furthermore, since $\langle \mathbf{x}, \mathbf{b}_j^* \rangle \in \mathbb{Z}$, $|\langle \mathbf{x}, \mathbf{b}_j^* \rangle| \geq 1$. Therefore

$$1 \leq |\langle \mathbf{x}, \mathbf{b}_j^* \rangle| \leq \|\mathbf{x}\|_2 \|\mathbf{b}_j^*\|_2 \Rightarrow \|\mathbf{x}\|_2 \geq \frac{1}{\|\mathbf{b}_j^*\|_2} \geq \min_{1 \leq i \leq k} \frac{1}{\|\mathbf{b}_i^*\|_2}$$

Since the final lower bound holds for all $\mathbf{x} \in G \setminus \{\mathbf{0}\}$, we get that $\lambda_1(G) > \min_{1 \leq i \leq k} \frac{1}{\|\mathbf{b}_i^*\|_2}$, as needed. Since $\lambda_1(G) > 0$, we get that G is a lattice. Lastly, note that G^* contains k linearly independent vectors if and only $\dim(G^*) = k = \dim(G)$. $\square$

We leave the following as an exercise.

Exercise 1 For any additive subgroup $G \subseteq \mathbb{R}^n$, show that G^* is a lattice.

Definition 3 (Pseudo-Inverse) *For a non-singular matix $B = (\mathbf{b}_1, \ldots, \mathbf{b}_k) \in \mathbb{R}^{n \times k}$, we define its pseudo-inverse to be $B^+ \stackrel{\text{def}}{=} (B^t B)^{-1} B^t \in \mathbb{R}^{k \times n}$. We define $\mathbf{b}_1^*, \ldots, \mathbf{b}_k^*$ to the columns of $(B^+)^t$, i.e. $(B^+)^t = (\mathbf{b}_1^*, \ldots, \mathbf{b}_k^*)$, and denote this set of vectors to be the associated dual basis for $\mathbf{b}_1, \ldots, \mathbf{b}_k$.*

Remark 4 Note that when B is a square matrix (and hence invertible), the pseudo-inverse corresponds to the standard inverse since $B^+ = (B^t B)^{-1} B^t = B^{-1} B^{-t} B^t = B^{-1}$.

In the next lemma, we establish the basic properties of the pseudo-inverse.

LEMMA 5 *Let $B = (\mathbf{b}_1, \ldots, \mathbf{b}_k) \in \mathbb{R}^{n \times k}$ be a non-singular matrix. The following holds:*

1. *B^+ is well-defined.*
2. *The dual basis vectors $\mathbf{b}_1^*, \ldots, \mathbf{b}_k^*$ are contained in $\text{span}(B)$ and are linearly independent.*
3. *$\ker(B^+) = \text{span}(B)^\perp$.*
4. *$B^+ B = I_k$, where I_k is the $k \times k$ identity.*
5. *For $i, j \in k$, $\langle \mathbf{b}_i^*, \mathbf{b}_j \rangle = \delta_{ij}$, where $\delta_{ij} = 1$ if $i = j$ and 0 otherwise.*

PROOF:

Proof of 1. Since $B^+ = (B^t B)^{-1} B^t$, to prove that B^+ is well-defined we need only show that $B^t B \in \mathbb{R}^{k \times k}$ is invertible. In particular, it suffices to show that $B^t B$ is non-singular. Take $\mathbf{x} \in \mathbb{R}^k$, and assume that $B^t B \mathbf{x} = \mathbf{0}$. We need to prove that $\mathbf{x} = \mathbf{0}$. Note that $0 = \mathbf{x}^t B^t B \mathbf{x} = \|B\mathbf{x}\|_2^2$. From here, we see that $\|B\mathbf{x}\|_2^2 = 0 \Leftrightarrow B\mathbf{x} = \mathbf{0} \Leftrightarrow \mathbf{x} = \mathbf{0}$ since B is non-singular. Therefore $\mathbf{x} = \mathbf{0}$ as needed.

Proof of 2. Note that by definition of $\mathbf{b}_1^*, \ldots, \mathbf{b}_k^*$, we have that $(\mathbf{b}_1^*, \ldots, \mathbf{b}_k^*) = ((B^t B)^{-1} B^t)^t = B(B^t B)^{-t} = B(B^t B)^{-1}$, where the last equality follows since $B^t B$ is symmetric (and hence has a symmetric inverse). Therefore $\mathbf{b}_i^* = B\left((B^t B)^{-1} \mathbf{e}_i\right) \in \text{span}(B)$ as needed. Furthermore, since both B and $(B^t B)^{-1}$ are non-singular, it follows that $(B^+)^t = B(B^t B)^{-1}$ is non-singular. Since $\mathbf{b}_1^*, \ldots, \mathbf{b}_k^*$ correspond to the columns of $(B^+)^t$, it follows that $\mathbf{b}_1^*, \ldots, \mathbf{b}_k^*$ are linearly independent.

Proof of 3. Since $(B^t B)^{-1}$ is non-singular, note that for $\mathbf{x} \in \mathbb{R}^n$ we have that $B^+ \mathbf{x} = (B^t B)^{-1} \left(B^t \mathbf{x}\right) = \mathbf{0} \Leftrightarrow B^t \mathbf{x} = \mathbf{0} \Leftrightarrow \mathbf{x} \in \text{span}(B)^\perp$. Therefore $\mathbf{x} \in \ker(B^+) \Leftrightarrow \mathbf{x} \in \text{span}(B)^\perp$ as needed.

Proof of 4. $B^+ B = (B^t B)^{-1}(B^t B) = I_k$ as needed.

Proof of 5. For $i, j \in [k]$, $\langle \mathbf{b}_i^*, \mathbf{b}_j \rangle = \langle (B^+)^t \mathbf{e}_i, B\mathbf{e}_j \rangle = \mathbf{e}_i^t B^+ B \mathbf{e}_j = \mathbf{e}_i^t I_k \mathbf{e}_j = \delta_{ij}$ as needed. □

LEMMA 6 *Let $\mathcal{L} \subseteq \mathbb{R}^n$ be a rank $k \geq 1$ lattice with basis $B = (\mathbf{b}_1, \ldots, \mathbf{b}_k) \in \mathbb{R}^{n \times k}$. Then $\mathcal{L}^*$ is a rank k lattice with basis $\mathbf{b}_1^*, \ldots, \mathbf{b}_k^*$. Furthermore, $(\mathcal{L}^*)^* = \mathcal{L}$.*

PROOF: We wish to prove that $\mathcal{L}(\mathbf{b}_1^*, \ldots, \mathbf{b}_k^*) = \mathcal{L}^*$. We show that $\mathcal{L}(\mathbf{b}_1^*, \ldots, \mathbf{b}_k^*) \subseteq \mathcal{L}^*$. First since $\mathbf{b}_1^*, \ldots, \mathbf{b}_k^* \in \text{span}(B) = \text{span}(\mathcal{L})$, we have that $\mathcal{L}(\mathbf{b}_1^*, \ldots, \mathbf{b}_k^*) \subseteq \text{span}(\mathcal{L})$. Now take $\mathbf{x} \in \mathcal{L}$ and $\mathbf{y} \in \mathcal{L}(\mathbf{b}_1^*, \ldots, \mathbf{b}_k^*)$. Since $\mathbf{b}_1, \ldots, \mathbf{b}_k$ is a basis for $\mathcal{L}$, we may express $\mathbf{x} = \sum_{i=1}^k a_i \mathbf{b}_i$ for $a_1, \ldots, a_k \in \mathbb{Z}$. Similarly, $\mathbf{y} = \sum_{i=1}^k b_i \mathbf{b}_i^*$ with $b_1, \ldots, b_k \in \mathbb{Z}$. Since $\langle \mathbf{b}_i, \mathbf{b}_j^* \rangle = \delta_{ij}$, $i, j \in [k]$, we have that

$$\langle \mathbf{x}, \mathbf{y} \rangle = \langle \sum_{i=1}^k a_i \mathbf{b}_i, \sum_{j=1}^k b_i \mathbf{b}_j^* \rangle = \sum_{1 \leq i,j \leq k} a_i b_j \langle \mathbf{b}_i, \mathbf{b}_j^* \rangle = \sum_{i=1}^k a_i b_i \in \mathbb{Z}$$

since $a_i, b_i \in \mathbb{Z}$, $i \in [k]$. Therefore $\mathcal{L}(\mathbf{b}_1^*, \ldots, \mathbf{b}_k^*) \subseteq \mathcal{L}^*$ as needed. We now prove that $\mathcal{L}^* \subseteq \mathcal{L}(\mathbf{b}_1^*, \ldots, \mathbf{b}_k^*)$. Take $\mathbf{y} \in \mathcal{L}^*$. Examine $\hat{\mathbf{y}} = \sum_{i=1}^k \langle \mathbf{y}, \mathbf{b}_i \rangle \mathbf{b}_i^*$. Since $\langle \mathbf{y}, \mathbf{b}_i \rangle \in \mathbb{Z}$, $i \in [k]$, we clearly have that $\hat{\mathbf{y}} \in \mathcal{L}(\mathbf{b}_1^*, \ldots, \mathbf{b}_k^*)$. For $j \in [k]$, note that

$$\langle \mathbf{y} - \hat{\mathbf{y}}, \mathbf{b}_j \rangle = \langle \mathbf{y}, \mathbf{b}_j \rangle - \langle \sum_{i=1}^k \langle \mathbf{y}, \mathbf{b}_i \rangle \mathbf{b}_i^*, \mathbf{b}_j \rangle = \langle \mathbf{y}, \mathbf{b}_j \rangle - \sum_{i=1}^k \langle \mathbf{y}, \mathbf{b}_i \rangle \delta_{ij} = \langle \mathbf{y}, \mathbf{b}_j \rangle - \langle \mathbf{y}, \mathbf{b}_j \rangle = 0.$$

From the above, we see that $\mathbf{y} - \hat{\mathbf{y}} \in \text{span}(B)^\perp$. Furthermore, since $\mathbf{y}, \hat{\mathbf{y}} \in \text{span}(B)$, we get that $\mathbf{y} - \hat{\mathbf{y}} \in \text{span}(B) \cap \text{span}(B)^\perp \Rightarrow \mathbf{y} - \hat{\mathbf{y}} = \mathbf{0}$. Therefore $\mathbf{y} = \hat{\mathbf{y}} \Rightarrow \mathbf{y} \in \mathcal{L}(\mathbf{b}_1^*, \ldots, \mathbf{b}_k^*)$ as needed.

For the furthermore, we wish to show that $(\mathcal{L}^*)^* = \mathcal{L}$. Since $\mathbf{b}_1^*, \ldots, \mathbf{b}_k^* \in \text{span}(B)$ and linearly independent, we clearly have that $\text{span}(\mathcal{L}) = \text{span}(B) = \text{span}(\mathcal{L}^*)$. Next, since for all $\mathbf{x} \in \mathcal{L}$, $\mathbf{y} \in \mathcal{L}^*$, we have that $\langle \mathbf{x}, \mathbf{y} \rangle \in \mathbb{Z}$, we get that $\mathcal{L} \subseteq (\mathcal{L}^*)^*$. Now take $\mathbf{z} \in (\mathcal{L}^*)^*$. Since $\mathbf{z} \in \text{span}(B)$, we can write $\mathbf{z} = \sum_{i=1}^k a_i \mathbf{b}_i$ for $a_1, \ldots, a_k \in \mathbb{R}$. Since $\mathbf{z} \in (\mathcal{L}^*)^*$, we note that $\langle \mathbf{z}, \mathbf{b}_i^* \rangle = a_i \in \mathbb{Z}$, for all $i \in [k]$. Therefore $\mathbf{z} \in \mathcal{L}$ as needed. □

Given the above lemmas, we get an alternate characterization of lattices.

THEOREM 7 *Let $G \subseteq \mathbb{R}^n$ be an additive subgroup. Then the following are equivalent:*

1. *G is a lattice.*
2. *$(G^*)^* = G$.*
3. *$\dim(G^*) = \dim(G)$.*

PROOF:

$1 \Rightarrow 2$. Follows directly from Lemma 6.

$2 \Rightarrow 3$. Since $G^* \subseteq \text{span}(G)$, we have the trivial inequality $\dim(G^*) \leq \dim(G)$. Since $(G^*)^* = G$, we have that $\dim(G) \geq \dim(G^*) \geq \dim((G^*)^*) = \dim(G)$. Therefore $\dim(G) = \dim(G^*)$ as needed.

$3 \Rightarrow 1$. Follows directly from Lemma 2. □

Using the dual lattice, we can also get an exact description of when the orthogonal projection of an lattice remains a lattice.

LEMMA 8 *Let $\mathcal{L} \subseteq \mathbb{R}^n$ be a lattice. Let $S = \text{span}(\mathcal{L})$ and let $W \subseteq S$ be a linear subspace.*

1. *$\pi_W(\mathcal{L})^* = \mathcal{L}^* \cap W$. Furthermore, this holds if $\mathcal{L}$ is any additive subgroup.*
2. *$\pi_W(\mathcal{L})$ is a lattice $\Leftrightarrow \dim(W \cap \mathcal{L}^*) = \dim(W \cap S) \Leftrightarrow \dim(W^\perp \cap \mathcal{L}) = \dim(W^\perp \cap S)$.*

PROOF:

Proof of 1. We first show that $\pi_W(\mathcal{L})^* \subseteq \mathcal{L}^* \cap W$. Take $\mathbf{x} \in \pi_W(\mathcal{L})^*$. First, we clearly have that $\mathbf{x} \in \text{span}(\pi_W(\mathcal{L})) = \text{span}(\pi_W(S)) = W \subseteq S$. Next note that for $\mathbf{y} \in \mathcal{L}$, since $\pi_W(\mathbf{x}) = \mathbf{x}$ we have that $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, \pi_W(\mathbf{y}) \rangle \in \mathbb{Z}$ since $\pi_W(\mathbf{y}) \in \pi_W(\mathcal{L})$. Therefore $\mathbf{x} \in \mathcal{L}^* \cap W$ as needed. Now we show that $\mathcal{L}^* \cap W \subseteq \pi_W(\mathcal{L})^*$. Take $\mathbf{x} \in \mathcal{L}^* \cap W$. Note that $\mathbf{x} \in \text{span}(\mathcal{L}^*) \cap W \subseteq S \cap W = W$. Take any $\mathbf{y} \in \pi_W(\mathcal{L})$, and let $\hat{\mathbf{y}} \in \mathcal{L}$ be any lifting of $\mathbf{y}$ satisfying $\pi_W(\hat{\mathbf{y}}) = \mathbf{y}$. Since $\pi_W(\mathbf{x}) = \mathbf{x}$, we have that $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, \pi_W(\hat{\mathbf{y}}) \rangle = \langle \mathbf{x}, \hat{\mathbf{y}} \rangle \in \mathbb{Z}$ since $\hat{\mathbf{y}} \in \mathcal{L}$. Therefore $\mathbf{x} \in \pi_W(\mathcal{L})^*$ as needed. Since the proof uses only the properties of additive subgroups, the characterization holds when $\mathcal{L}$ is any additive subgroup.

Proof of 2. From Theorem 7 we have that $\pi_W(\mathcal{L})$ is a lattice $\Leftrightarrow \dim(\pi_W(\mathcal{L})^*) = \dim(\pi_W(\mathcal{L}))$. By the first part of the Lemma, we have that $\pi_W(\mathcal{L})^* = \mathcal{L}^* \cap W$, and hence $\dim(\pi_W(\mathcal{L})) = \dim(\pi_W(\mathcal{L})^*) \Leftrightarrow \dim(\pi_W(\mathcal{L})) = \dim(\mathcal{L}^* \cap W)$. Next, note that $\text{span}(\pi_W(\mathcal{L})) = \pi_W(\text{span}(\mathcal{L})) = W$. This proves that $\pi_W(\mathcal{L})$ is a lattice $\Leftrightarrow \dim(W) = \dim(W \cap S) = \dim(\mathcal{L}^* \cap W)$.

We now show that $\dim(W \cap S) = \dim(\mathcal{L}^* \cap W) \Leftrightarrow \dim(W^\perp \cap S) = \dim(\mathcal{L} \cap W^\perp)$. Since the statements are symmetric (i.e. since $(W^\perp)^\perp = W$ and $(\mathcal{L}^*)^* = \mathcal{L}$), it suffices to prove the implication one way. We assume that $\dim(W \cap S) = \dim(\mathcal{L}^* \cap W)$. Let $l = \dim(W \cap S)$ and let $k = \dim(S)$. Since $\mathcal{L} \cap W^\perp \subseteq S \cap W^\perp$, we clearly have that $\dim(\mathcal{L} \cap W^\perp) \leq \dim(S \cap W^\perp)$. It therefore suffices to prove the reverse inequality. By the assumption on $\mathcal{L}^* \cap W$, i.e. that $\dim(\mathcal{L}^* \cap W) = l$, we have that $\mathcal{L}^* \cap W$ has a basis $\mathbf{b}_1^*, \ldots, \mathbf{b}_l^* \in \mathcal{L}^* \cap W$. By Theorem 5 of Lecture 1, we can extend $\mathbf{b}_1^*, \ldots, \mathbf{b}_l^*$ to a basis $\mathbf{b}_1^*, \ldots, \mathbf{b}_k^*$ of $\mathcal{L}^*$. Let $\mathbf{b}_1, \ldots, \mathbf{b}_k$ denote the associated basis of $\mathcal{L}$, i.e. which satisfies $\langle \mathbf{b}_i, \mathbf{b}_j^* \rangle = \delta_{ij}$, for $i, j \in [k]$. Note that $\langle \mathbf{b}_i^*, \mathbf{b}_j \rangle = 0$ for $i \in \{1, \ldots, l\}$ and $j \in \{l+1, \ldots, k\}$. Therefore $\mathbf{b}_{l+1}, \ldots, \mathbf{b}_k \in \text{span}(\mathcal{L}) \cap \text{span}(\mathcal{L}^* \cap W)^\perp = S \cap (S \cap W)^\perp = S \cap W^\perp$. Since $\mathbf{b}_{l+1}, \ldots, \mathbf{b}_k$ are linearly independent, we have that

$$\dim(\mathcal{L} \cap W^\perp) \geq \dim(\text{span}(\mathbf{b}_{l+1}, \ldots, \mathbf{b}_k)) = k - l = \dim(S) - \dim(S \cap W) = \dim(S \cap W^\perp)$$

as needed. □

Using the previous lemma, we easily derive the following corollary.

Corollary 9 *Let $\mathcal{L} \subseteq \mathbb{R}^n$ be a lattice.*

1. *For linearly independent vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathcal{L}$, let π denote the orthogonal projection onto $\text{span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)^\perp$. Then $\pi(\mathcal{L})$ is a lattice.*
2. *For linearly independent vectors $\mathbf{v}_1^*, \ldots, \mathbf{v}_k^* \in \mathcal{L}^*$, let π denote the orthogonal projection onto $\text{span}(\mathbf{v}_k^*, \ldots, \mathbf{v}_k^*)$. Then $\pi(\mathcal{L})$ is a lattice.*

Proof:

Proof of 1. Let $W = \text{span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)^\perp \cap \text{span}(\mathcal{L})$. Since $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathcal{L}$ note that $\pi(\mathcal{L}) = \pi_W(\mathcal{L})$. Since $W^\perp = \text{span}(\mathbf{v}_1, \ldots, \mathbf{v}_k) + \text{span}(\mathcal{L})^\perp$ and $\mathbf{v}_1, \ldots, \mathbf{b}_k \in \mathcal{L}$, we see that $W^\perp \cap \text{span}(\mathcal{L}) = \text{span}(\mathbf{v}_1, \ldots, \mathbf{v}_k) = \text{span}(\mathcal{L} \cap W^\perp)$ and hence $\dim(\text{span}(\mathcal{L}) \cap W^\perp) = \dim(\mathcal{L} \cap W^\perp)$. Therefore by Lemma 8, we have that $\pi_W(\mathcal{L}) = \pi(\mathcal{L})$ is a lattice as needed.

Proof of 2. Let $W = \text{span}(\mathbf{v}_1^*, \ldots, \mathbf{v}_k^*)$. Given that $\mathcal{L}$ is a lattice, we know that $\text{span}(\mathcal{L}) = \text{span}(\mathcal{L}^*)$. Since $\mathbf{v}_1^*, \ldots, \mathbf{v}_k^* \in \mathcal{L}^*$, we see that $W \subseteq \text{span}(\mathcal{L}^*) = \text{span}(\mathcal{L})$ and hence $W \cap \text{span}(\mathcal{L}) = W = \text{span}(\mathcal{L}^* \cap W)$. Therefore $\dim(W \cap \text{span}(\mathcal{L})) = \dim(W \cap \mathcal{L}^*)$. By Lemma 8, we now have that $\pi_W(\mathcal{L}) = \pi(\mathcal{L})$ is a lattice as needed. □

2 Deciding Lattice Membership and Building a SubLattice Basis

In this section, we will focus on solving some "easy", but useful, computational problems on lattices. The two main tasks we will address are the following: let $\mathcal{L} \subseteq \mathbb{R}^n$ be a rank k lattice with basis $B \in \mathbb{R}^{n \times k}$. Let $\mathbf{y}_1, \ldots, \mathbf{y}_m$ be vectors in $\mathcal{L}$ (not necessarily linearly independent).

1. **Lattice Basis Problem.** Determine a basis for $\mathcal{L}(\mathbf{y}_1, \ldots, \mathbf{y}_m)$.

2. **Lattice Membership Problem.** Given $\mathbf{x} \in \mathbb{R}^n$, decide whether $\mathbf{x} \in \mathcal{L}(\mathbf{y}_1, \ldots, \mathbf{y}_m)$? If so, find an integer combination $z_1, \ldots, z_n \in \mathbb{Z}$ such that $\mathbf{x} = \sum_{i=1}^{m} z_i \mathbf{b}_i$.

REMARK 10 For the first question, since $\mathcal{L}(\mathbf{y}_1, \ldots, \mathbf{y}_m)$ is an additive subgroup of $\mathcal{L}$, it is clearly discrete and hence a lattice. Therefore it makes sense to ask for a basis of $\mathcal{L}(\mathbf{y}_1, \ldots, \mathbf{y}_m)$.

Simplications and Reductions. We now make some direct simplications. Letting $B^+ \in \mathbb{R}^{k \times n}$, denote the pseudo-inverse of B, we note that finding a basis of $\mathcal{L}(\mathbf{b}_1, \ldots, \mathbf{b}_m)$ is equivalent to finding a basis for $\mathcal{L}(B^+\mathbf{y}_1, \ldots, B^+\mathbf{y}_m)$. Note that each $B^+\mathbf{y}_i \in \mathbb{Z}^k$, $i \in [m]$, since B^+ gives the coordinates of $\mathbf{y}_i$ in terms of B. To see the equivalence, note that given a basis $\mathbf{z}_1, \ldots, \mathbf{z}_l \in \mathbb{Z}^k$ for $\mathcal{L}(B^+\mathbf{y}_1, \ldots, B^+\mathbf{y}_m) \subseteq \mathbb{Z}^k$, we get that $B\mathbf{z}_1, \ldots, B\mathbf{z}_l$ is a basis for $\mathcal{L}(\mathbf{y}_1, \ldots, \mathbf{y}_m)$. Therefore, we may assume that all the given vectors are integer vectors.

For the second question, we can make analoguous simplifications. First, using linear algebra, one may directly check whether $\mathbf{x} \in \text{span}(\mathbf{y}_1, \ldots, \mathbf{y}_m)$. Next, we may reduce to a problem on integer vectors as above, where we now check whether $B^+\mathbf{x} \in \mathcal{L}(B^+\mathbf{y}_1, \ldots, B^+\mathbf{y}_m) \subseteq \mathbb{Z}^k$. Clearly if $B^+\mathbf{x} \notin \mathbb{Z}^k$, then $B^+\mathbf{x} \notin \mathcal{L}(\mathbf{y}_1, \ldots, \mathbf{y}_m)$, and hence we may assume that $B^+\mathbf{x} \in \mathbb{Z}^k$ as well.

Thus far, we have reduced both questions 1 and 2 to the case where all the vectors are integral. We now show that the decisional version of question 2 reduces to question 1. Given $\mathbf{y}_1, \ldots, \mathbf{y}_m \in \mathbb{Z}^n$, let $\mathbf{z}_1, \ldots, \mathbf{z}_k \in \mathbb{Z}^k$ denote a basis for $\mathcal{L}(\mathbf{y}_1, \ldots, \mathbf{y}_m)$ as produced by any basis finding algorithm (solver for question 1). We now reduce checking whether $\mathbf{x} \in \mathcal{L}(\mathbf{y}_1, \ldots, \mathbf{y}_m) = \mathcal{L}(\mathbf{z}_1, \ldots, \mathbf{z}_k)$ to solving the linear system of equations $(\mathbf{z}_1, \ldots, \mathbf{z}_k)\mathbf{a} = \mathbf{x}$ for $\mathbf{a} \in \mathbb{R}^k$. Here we have three cases, either (1) the system no solution, (2) the system has a non-integer solution, or (3) the system has an integer solution. Since $\mathbf{z}_1, \ldots, \mathbf{z}_k$ is a basis for the lattice, note that both cases (1) and (2) imply that $\mathbf{x} \notin \mathcal{L}(\mathbf{y}_1, \ldots, \mathbf{y}_m)$, and case (3) implies that $\mathbf{x} \in \mathcal{L}(\mathbf{y}_1, \ldots, \mathbf{y}_m)$, as needed. This completes the reduction. We note that in the case that $\mathbf{x} \in \mathcal{L}(\mathbf{y}_1, \ldots, \mathbf{y}_m)$ the previous reduction does not give specific integer multipliers $a_1, \ldots, a_m$ such that $\mathbf{x} = \sum_{i=1}^{m} a_i \mathbf{y}_i$. However, as we shall see, the techniques used to generate the basis for $\mathcal{L}(\mathbf{y}_1, \ldots, \mathbf{y}_k)$ will yield a method to compute these coefficients.

2.1 Applications

In the following section, we give two direct applications of the questions from the previous section.

Solving Integer Linear Systems of Equations. Given $A \in \mathbb{Z}^{n \times m}$, $\mathbf{b} \in \mathbb{Z}^n$, decide whether the system

$$A\mathbf{x} = \mathbf{b}, \mathbf{x} \in \mathbb{Z}^m, \text{has a solution}, \tag{1}$$

and if so, find a satisfying $\mathbf{x} \in \mathbb{Z}^m$.

PROPOSITION 11 *The Integer Linear System problem is equivalent to the Lattice Membership Problem.*

PROOF: Let $\mathcal{L} = \mathcal{L}(A)$ denote the integer sublattice generated by the columns of A. By definition, $A\mathbf{x} = \mathbf{b}$, $\mathbf{x} \in \mathbb{Z}^m$ has a solution $\Leftrightarrow$ $\mathbf{b}$ is an integral combinations of the columns of $A \Leftrightarrow \mathbf{b} \in \mathcal{L}(A)$. Hence the integer linear system problem is equivalent to the lattice membership problem

for integer sublattices. However, from the remarks in the previous section, we know that the lattice membership problem reduces to the case where all the vectors are integral, and hence the problems are completely equivalent. □

REMARK 12 We note that the when the columns of A are linearly independent, the above problem reduces to linear algebra. This is exactly the setting where the columns of A form a basis of $\mathcal{L}(A)$.

Solving Modular Systems of Equations. Given $A \in \mathbb{Z}^{n\times m}$, $\mathbf{b} \in \mathbb{Z}^n$, $\mathbf{c} \in \mathbb{Z}^n_+$, decide whether the system

$$A\mathbf{x} = \mathbf{b}, \mathbf{x} \in \mathbb{Z}^m A\mathbf{x} \equiv \begin{pmatrix} b_1 & (\text{mod } c_1) \\ & \vdots \\ b_n & (\text{mod } c_n) \end{pmatrix}, \mathbf{x} \in \mathbb{Z}^m, \text{ has a solution,} \tag{2}$$

and if so, find a satisfying $\mathbf{x} \in \mathbb{Z}^m$.

A natural question is when do integer linear systems of equations occur in practice where the columns of A are not linearly independent? As we will see in the following reduction, the problem of solving modular equations reduces to solving an integer linear system where we do not have linear independence.

PROPOSITION 13 *Solving a Modular System of Equations reduces to solving a Integer Linear System.*

PROOF: By definition, we have that $a \equiv b \pmod{c}$, $a, b \in \mathbb{Z}$, $c \in \mathbb{N}$, if and only if $a + zc = b$ for some $z \in \mathbb{Z}$. By the identical reasoning, we have that

$$A\mathbf{x} \equiv \begin{pmatrix} b_1 & (\text{mod } c_1) \\ & \vdots \\ b_n & (\text{mod } c_n) \end{pmatrix}, \mathbf{x} \in \mathbb{Z}^m, \text{ has a solution} \Leftrightarrow$$

$$A\mathbf{x} + \begin{pmatrix} c_1 & 0 & \dots & 0 \\ 0 & c_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & c_n \end{pmatrix} \mathbf{z} = \mathbf{b}, \mathbf{x} \in \mathbb{Z}^m, \mathbf{z} \in \mathbb{Z}^n, \text{ has a solution.}$$

Furthermore, given a solution $(\mathbf{x}, \mathbf{z})$ to the integer linear system, the vector $\mathbf{x}$ gives a solution to the modular system. Hence, the modular system of equations reduces to an integer system of equations. Lastly, note that the columns of the extended coefficient matrix are *n*ot linearly independent (since the diagonal matrix of moduli spans $\mathbb{R}^n$). □

The following exercise shows that both of the above problems either have solutions or have succinct proofs of infeasibility.

EXERCISE 2 (DUALITY FOR INTEGER LINEAR SYSTEMS) Take $A \in \mathbb{Z}^{n\times m}$, $\mathbf{b} \in \mathbb{Z}^n$, $\mathbf{c} \in \mathbb{Z}^n_+$.

1. Prove that the system $A\mathbf{x} = \mathbf{b}$, $\mathbf{x} \in \mathbb{Z}^m$ has a solution if and only if there does not exist $\mathbf{y} \in \mathbb{R}^n$ such that $\mathbf{y}^t A \in \mathbb{Z}^m$ and $\mathbf{y}^t\mathbf{b} \notin \mathbb{Z}$. (Hint: split up the analysis based on whether $A\mathbf{x} = \mathbf{b}$ has a real solution or not. If it has a real solution, examine the appropriate dual lattice.)

2. Prove that the system

$$A\mathbf{x} \equiv \begin{pmatrix} b_1 & (\text{mod } c_1) \\ & \vdots \\ b_n & (\text{mod } c_n) \end{pmatrix}, \mathbf{x} \in \mathbb{Z}^m, \text{ has a solution,}$$

if and only if there does not exist $\mathbf{y} \in \mathbb{R}^n$, $y_i \in \{0, \frac{1}{c_i}, \ldots, \frac{c_i-1}{c_i}\}$, $i \in [n]$, such that $\mathbf{y}^t A \in \mathbb{Z}^m$ and $\mathbf{y}^t \mathbf{b} \notin \mathbb{Z}$.

2.2 Hermite Normal Form

In this section, we describe a method for solving the lattice basis problem for finitely generated integer sublattices. As mentioned previously, this suffices to solve the general lattice basis problem, as well as the decisional version of the lattice membership problem.

Definition 14 (Hermite Normal Form) *Let $A \in \mathbb{Z}^{n \times m}$, and let $k = \text{rank}(A)$. Let $r(i)$, $i \in [m]$, denote the index of the first non-zero entry in the i^{th} column of A.*
The matrix A is in Hermite Normal Form (HNF) if it satisfies the following:

1. *The first k columns of A are non-zero, and the remaining are zero.*
2. *$r(1) < r(2) < \cdots < r(k)$.*
3. *$A_{r(i),i} > 0$ for all $i \in [k]$.*
4. *$0 \leq A_{r(i),j} < A_{r(i),i}$ for all $1 \leq j < i \leq k$.*

The main goal of this section is to prove the following theorem:

Theorem 15 *For any $A \in \mathbb{Z}^{n \times m}$, there exists a unimodular transformation $U \in \mathbb{Z}^{m \times m}$ such that AU is in HNF. Furthermore, if AU_1 and AU_2 are in HNF, $U_1, U_2 \in \mathbb{Z}^{m \times m}$ unimodular, then $AU_1 = AU_2$.*

The above theorem is constructive: we will give an algorithm which computes both the HNF and the corresponding unimodular transformation for any integer matrix $A \in \mathbb{Z}^{n \times m}$. Using the above procedure for computing the HNF, the following algorithm computes a basis for $\mathcal{L}(A)$:

1. Compute the HNF $\bar{A} = AU$ for A, $U \in \mathbb{Z}^{m \times m}$ unimodular.
2. Return the first $\text{rank}(\bar{A})$ columns of $\bar{A}$.

We now justify the correctness of the above procedure. Let $k = \text{rank}(\bar{A})$ and let $\bar{A}_{\cdot,[k]}$ denote the first k columns of $\bar{A}$. Since the remaining columns of $\bar{A}$ are zero, we clearly have that $\mathcal{L}(\bar{A}_{\cdot,[k]}) = \mathcal{L}(\bar{A})$. Next, since $\bar{A} = AU$ for a unimodular $U \in \mathbb{Z}^{m \times m}$, we have that $\mathcal{L}(\bar{A}) = \mathcal{L}(AU) = \mathcal{L}(A)$. Furthermore, note that $k = \text{rank}(\bar{A}) = \text{rank}(A)$, since U is non-singular. Given this we have that the columns of $\bar{A}_{\cdot,[k]}$ are linearly independent and that $\mathcal{L}(\bar{A}_{\cdot,[k]}) = \mathcal{L}(A)$. Hence $\bar{A}_{\cdot,[k]}$ form a basis of $\mathcal{L}(A)$ as desired.

To construct the HNF we will use the following elementary integer columns operations:

1. $A_{\cdot,i} \leftarrow A_{\cdot,i} + zA_{\cdot,j}$, $z \in \mathbb{Z}$, $i \neq j$. (add integer multiple of column j to column i)
2. $A_{\cdot,i} \leftrightarrow A_{\cdot,j}$, $i \neq j$. (swap column i and j)

3. $A_{\cdot,i} \leftarrow -A_{\cdot,i}$. (negate column i)

We leave it as an exercise to show that each of the above operations corresponds to multiplying the matrix A on the right by a unimodular transformation. Furthermore, any sequence of such operations also corresponds to right mutliplication by a unimodular, since unimodular transformations form a group under multiplication.

PROOF:[of Theorem 15]

Given a matrix $A \in \mathbb{Z}^{m \times n}$, we shall use the following algorithm to put it in HNF:

Require: $A \in \mathbb{Z}^{n \times m}$
Ensure: A in HNF.
  $c \leftarrow 0$ { lower bound on rank(A) }
  for all $r \in 1$ **to** n **do**
    if $A_{r,[c+1,m]} \neq \mathbf{0}$ **then**
      $c \leftarrow c + 1$ { increase rank lower bound by 1}
      for all $i \in c$ **to** m **do**
        $A_{\cdot,i} \leftarrow \text{sign}(A_{r,i}) \cdot A_{\cdot,i}$ { make partial row non-negative }
      repeat
        $i \leftarrow \arg\min\{A_{r,j} : A_{r,j} \neq 0, c \leq j \leq m\}$ { find index of smallest non-zero entry }
        $A_{\cdot,c} \leftrightarrow A_{\cdot,i}$ { swap column c with smallest non-zero entry column }
        for all $j \in c + 1$ **to** m **do**
          $A_{\cdot,j} \leftarrow A_{\cdot,j} - \lfloor \frac{A_{r,j}}{A_{r,c}} \rfloor A_{\cdot,c}$ { main Euclidean algorithm step }
      until $A_{r,[c+1,m]} = \mathbf{0}$
      for all $j \in 1$ **to** $c - 1$ **do**
        $A_{\cdot,j} \leftarrow A_{\cdot,j} - \lfloor \frac{A_{r,j}}{A_{r,c}} \rfloor A_{\cdot,c}$ { row cleanup }

We shall prove that the above algorithm terminates in a bounded number of iterations and that it puts A in HNF. □

We note that the algorithm presented above is not known to terminate in polynomial time. The main issue is that the as written, the size of the intermediate numbers in the working matrix A could be very very large (i.e. have a super-polynomial number of bits). There are many ways to overcome this problem, though we shall not cover them here. For a thorough reference on the Hermite Normal Form, one may consult [Sch86].

References

[Sch86] A. Schrijver. *Theory of Linear and Integer Programming*. Wiley-Interscience, New York, NY, 1986.