CSE 206A: Lattice Algorithms and Applications	Winter 2010
1: Introduction to Lattices	
Instructor: *Daniele Micciancio*	UCSD CSE

Lattices are regular arrangements of points in Euclidean space. They naturally occur in many settings, like crystallography, sphere packings (stacking oranges), etc. They have many applications in computer science and mathematics, including the solution of integer programming problems, diophantine approximation, cryptanalysis, the design of error correcting codes for multi antenna systems, and many more. Recently, lattices have also attracted much attention as a source of computational hardness for the design of secure cryptographic functions.

This course offers an introduction to lattices. We will study the best currently known algorithms to solve the most important lattice problems, and how lattices are used in several representative applications. We begin with the definition of lattices and their most important mathematical properties.

1. Lattices

Definition 1. A *lattice* is a *discrete* additive *subgroup* of $\mathbb{R}^n$, i.e., it is a subset $\Lambda \subseteq \mathbb{R}^n$ satisfying the following properties:

(subgroup) Λ is closed under addition and subtraction,[1]

(discrete) there is an $\epsilon > 0$ such that any two distinct lattice points $\mathbf{x} \neq \mathbf{y} \in \Lambda$ are at distance at least $\|\mathbf{x} - \mathbf{y}\| \geq \epsilon$.

Not every subgroup of $\mathbb{R}^n$ is a lattice.

Example 1. $\mathbb{Q}^n$ is a subgroup of $\mathbb{R}^n$, but not a lattice, because it is not discrete.

The simplest example of lattice is the set of all n-dimensional vectors with integer entries.

Example 2. The set $\mathbb{Z}^n$ is a lattice because integer vectors can be added and subtracted, and clearly the distance between any two integer vectors is at least 1.

Other lattices can be obtained from $\mathbb{Z}^n$ by applying a (nonsingular) linear transformation. For example, if $\mathbf{B} \in \mathbb{R}^{k\times n}$ has full column rank (i.e., the columns of $\mathbf{B}$ are linearly independent), then $\mathbf{B}(\mathbb{Z}^n) = \{\mathbf{Bx} \colon \mathbf{x} \in \mathbb{Z}^n\}$ is also a lattice. Clearly this set is closed under addition and subtraction. Later we will show that it is also discrete, and moreover all lattices can be expressed as $\mathbf{B}(\mathbb{Z}^n)$ for some $\mathbf{B}$, so an equivalent definition of lattice is the following.

Definition 2. Let $\mathbf{B} = [\mathbf{b}_1, \ldots, \mathbf{b}_k] \in \mathbb{R}^{n\times k}$ be linearly independent vectors in $\mathbb{R}^n$. The lattice generated by $\mathbf{B}$ is the set

$$\mathcal{L}(\mathbf{B}) = \left\{\mathbf{Bx} \colon \mathbf{x} \in \mathbb{Z}^k\right\} = \left\{\sum_{i=1}^{k} x_i \cdot \mathbf{b}_i \colon x_i \in \mathbb{Z}\right\}$$

of all the *integer* linear combinations of the columns of $\mathbf{B}$. The matrix $\mathbf{B}$ is called a *basis* for the lattice $\mathcal{L}(\mathbf{B})$. The integers n and k are called the *dimension* and *rank* of the lattice. If $n = k$ then $\mathcal{L}(\mathbf{B})$ is called a *full rank* lattice.

---

[1] Technically, closure under subtraction is enough because addition can be expressed as $a + b = a - (-b)$.

Definition 2 is the most commonly used in comptuer science as it gives a natural way to represent a lattice by a finite object: lattices are represented by a basis matrix $\mathbf{B}$ that generates the lattice, and the basis matrix typically has integer or rational entries.

Notice the similarity between the definition of a lattice

$$\mathcal{L}(\mathbf{B}) = \{\mathbf{B} \cdot \mathbf{x} : \mathbf{x} \in \mathbb{Z}^k\}.$$

and the definition of vector space generated by $\mathbf{B}$:

$$\text{span}(\mathbf{B}) = \{\mathbf{B} \cdot \mathbf{x} \colon \mathbf{x} \in \mathbb{R}^k\}.$$

The difference is that in a vector space you can combine the columns of $\mathbf{B}$ with arbitrary real coefficients, while in a lattice only integer coefficients are allowed, resulting in a discrete set of points. Notice that, since vectors $\mathbf{b}_1, \ldots, \mathbf{b}_n$ are linearly independent, any point $\mathbf{y} \in \text{span}(\mathbf{B})$ can be written as a linear combination $\mathbf{y} = x_1\mathbf{b}_1 + \cdots x_n\mathbf{b}_n$ in a unique way. Therefore $\mathbf{y} \in \mathcal{L}(\mathbf{B})$ if and only if $x_1, \ldots, x_n \in \mathbb{Z}$.

If $\mathbf{B}$ is a basis for the lattice $\mathcal{L}(\mathbf{B})$, then it is also a basis for the vector space $\text{span}(\mathbf{B})$. However, not every basis for the vector space $\text{span}(\mathbf{B})$ is also a lattice basis for $\mathcal{L}(\mathbf{B})$. For example $2\mathbf{B}$ is a basis for $\text{span}(\mathbf{B})$ as a vector space, but it is not a basis for $\mathcal{L}(\mathbf{B})$ as a lattice because vector $\mathbf{b}_i \in \mathcal{L}(\mathbf{B})$ (for any i) is not an *integer* linear combination of the vectors in $2\mathbf{B}$.

The definition $\mathcal{L}(\mathbf{B}) = \{\mathbf{B}\mathbf{x} \colon \mathbf{x} \in \mathbb{Z}^m\}$ can be extended to matrices $\mathbf{B}$ whose columns are not linearly independent. However, in this case, the resulting set of points is not always a lattice because it may not be discrete. Still, we will see that if $\mathbf{B}$ is a matrix with rational entries, then $\mathcal{L}(\mathbf{B})$ is always a lattice, and a basis for $\mathcal{L}(\mathbf{B})$ can be computed from $\mathbf{B}$ in polynomial time.

Exercise 1. Find a set of vecotrs $\mathbf{B}$ such that $\mathcal{L}(\mathbf{B})$ is not a lattice. *[Hint: these vectors must necessarily be linearly dependent and irrational.]*

2. Lattice bases

Theorem 1. *Let $\mathbf{B}$ and $\mathbf{C}$ be two bases. Then $\mathcal{L}(\mathbf{B}) = \mathcal{L}(\mathbf{C})$ if and only if there exists a unimodular matrix $\mathbf{U}$ (i.e., a square matrix with integer entries and determinant ± 1) such that $\mathbf{B} = \mathbf{C}\mathbf{U}$.*

Proof. First assume $\mathbf{B} = \mathbf{C}\mathbf{U}$ for some unimodular matrix $\mathbf{U}$. Notice that if $\mathbf{U}$ is unimodular, then $\mathbf{U}^{-1}$ is also unimodular. In particular, both $\mathbf{U}$ and $\mathbf{U}^{-1}$ are integer matrices, and $\mathbf{B} = \mathbf{C}\mathbf{U}$ and $\mathbf{C} = \mathbf{B}\mathbf{U}^{-1}$. It follows that $\mathcal{L}(\mathbf{B}) \subseteq \mathcal{L}(\mathbf{C})$ and $\mathcal{L}(\mathbf{C}) \subseteq \mathcal{L}(\mathbf{B})$, i.e., the two matrices $\mathbf{B}$ and $\mathbf{C}$ generate the same lattice.

Now assume $\mathbf{B}$ and $\mathbf{C}$ are two bases for the same lattice $\mathcal{L}(\mathbf{B}) = \mathcal{L}(\mathbf{C})$. Then, by definition of lattice, there exist integer square matrices $\mathbf{V}$ and $\mathbf{W}$ such that $\mathbf{B} = \mathbf{C}\mathbf{W}$ and $\mathbf{C} = \mathbf{B}\mathbf{V}$. Combining these two equations we get $\mathbf{B} = \mathbf{B}\mathbf{V}\mathbf{W}$, or equivalently, $\mathbf{B}(\mathbf{I} - \mathbf{V}\mathbf{W}) = \mathbf{O}$. Since vectors $\mathbf{B}$ are linearly independent, it must be $\mathbf{I} - \mathbf{V}\mathbf{W} = \mathbf{O}$, i.e., $\mathbf{V}\mathbf{W} = \mathbf{I}$. In particular, $\det(\mathbf{V}) \cdot \det(\mathbf{W}) = \det(\mathbf{V} \cdot \mathbf{W}) = \det(\mathbf{I}) = 1$. Since matrices $\mathbf{V}$ and $\mathbf{W}$ have integer entries, $\det(\mathbf{V}), \det(\mathbf{W}) \in \mathbb{Z}$, and it must be $\det(\mathbf{V}) = \det(\mathbf{W}) = \pm 1$ □

A simple way to obtain a basis of a lattice from another is to apply (a sequence of) elementary column operations, as defined below. It is easy to see that elementary column operations do not change the lattice generated by the basis because they can be expressed as right multiplication by a unimodular matrix. Elementary (integer) column operations are:

(1) Swap the order of two columns in $\mathbf{B}$.
(2) Multiply a column by -1.
(3) Add an integer multiple of a column to another column: $\mathbf{b}_i \leftarrow \mathbf{b}_i + a \cdot \mathbf{b}_j$ where $i \neq j$ and $a \in \mathbb{Z}$.

Moreover, any unimodular transformation can be expressed as a sequence of elementary integer column operations.

Exercise 2. Show that two lattice bases are equivalent precisely when one can be obtained from the other via elementary integer column operations. *[Hint: show that any unimodular matrix can be transformed into the identity matrix using elementary operations, and then reverse the sequence of operations.]*

3. Gram-Schmidt orthogonalization

Any basis $\mathbf{B}$ can be transformed into an orthogonal basis for the same vector space using the well-known Gram-Schmidt orthogonalization method. Suppose we have vectors $\mathbf{B} = [\mathbf{b}_1|\dots|\mathbf{b}_n] \in \mathbb{R}^{m\times n}$ generating a vector space $V = \mathrm{span}(\mathbf{B})$. These vectors are not necessarily orthogonal (or even linearly independent), but we can always find an orthogonal basis $\mathbf{B}^* = [\mathbf{b}_1^*|\dots|\mathbf{b}_n^*]$ for V where $\mathbf{b}_i^*$ is the component of $\mathbf{b}_i$ orthogonal to $\mathrm{span}(\mathbf{b}_1,\dots,\mathbf{b}_{i-1})$.

Definition 3. For any sequence of vectors $\mathbf{B} = [\mathbf{b}_1,\dots,\mathbf{b}_n]$, define the orthogonalized vectors $\mathbf{B}^* = [\mathbf{b}_1^*|\dots|\mathbf{b}_n^*]$ iteratively according to the formula

$$\mathbf{b}_i^* = \mathbf{b}_i - \sum_{j<i} \mu_{i,j}\mathbf{b}_j^* \text{ where } \mu_{i,j} = \frac{\langle \mathbf{b}_i, \mathbf{b}_j^*\rangle}{\langle \mathbf{b}_j^*, \mathbf{b}_j^*\rangle}.$$

In matrix notation, $\mathbf{B} = \mathbf{B}^*\mathbf{M}$ where $\mathbf{M}$ is the upper triangular matrix with 1 along the diagonal and $m_{j,i} = \mu_{i,j}$ for all $j < i$. It also follows that $\mathbf{B}^* = \mathbf{B}\mathbf{M}^{-1}$ where $\mathbf{M}^{-1}$ is also upper triangular with 1 along the diagonal.

Note that the columns of $\mathbf{B}^*$ are orthogonal ($\langle \mathbf{b}_i^*, \mathbf{b}_j^*\rangle = 0$ for all $i \neq j$). Therefore the (non-zero) columns of $\mathbf{B}^*$ are linearly independent and form a basis for the vector space $\mathrm{span}(\mathbf{B})$. However they are generally not a basis for the lattice $\mathcal{L}(\mathbf{B})$.

Example 3. The Gram-Schmidt orthogonalization of the basis $\mathbf{B} = [(2,0)^\top, (1,2)^\top]$ is $\mathbf{B}^* = [(2,0)^\top, (0,2)^\top]$. However this is not a lattice basis for $\mathcal{L}(\mathbf{B})$ because the vector $(0,2)^T$ does not belong to the lattice. $\mathcal{L}(\mathbf{B})$ contains a sublattice generated by a pair of orthogonal vectors $(2,0)^\top$ and $(0,4)^\top$, but no pair of orthogonal vectors generate the entire lattice $\mathcal{L}(\mathbf{B})$.

So, while vector spaces always admit an orthogonal basis, this is not true for lattices.

4. The determinant

Definition 4. Given a basis $\mathbf{B} = [\mathbf{b_1},\dots,\mathbf{b_n}] \in \mathbb{R}^{k\times n}$, the *fundamental parallelepiped* associated to $\mathbf{B}$ is the set of points

$$\mathcal{P}(\mathbf{B}) = \mathbf{B}[0,1)^n = \left\{\Sigma_{i=1}^n x_i \cdot \mathbf{b_i} \colon 0 \le x_i < 1\right\}.$$

Remark 1. Note that $\mathcal{P}(\mathbf{B})$ is half-open, so that the translates $\mathcal{P}(\mathbf{B}) + \mathbf{v}$ (for $\mathbf{v} \in \mathcal{L}(\mathbf{B})$) form a partition of the whole space $\mathbb{R}^k$. More precisely, for any $\mathbf{x} \in \mathbb{R}^m$, there exists a unique lattice point $\mathbf{v} \in \mathcal{L}(\mathbf{B})$, such that $\mathbf{x} \in (\mathbf{v} + \mathcal{P}(\mathbf{B}))$.

We now define a fundamental quantity associated to any lattice, the determinant.

Definition 5. Let $\mathbf{B} \in \mathbb{R}^{k \times n}$ be a basis. The determinant of a lattice $\det(\mathcal{L}(\mathbf{B}))$ is defined as the n-dimensional volume of the fundamental parallelepiped associated to $\mathbf{B}$:

$$\det(\mathcal{L}(B)) = \mathrm{vol}(\mathcal{P}(\mathbf{B})) = \prod_i \|\mathbf{b}_i^*\|$$

where $\mathbf{B}^*$ is the Gram-Schmidt orthogonalization of $\mathbf{B}$.

The above formula for the determinant of a lattice is a generalization of the well known formula for the area of a parallelepiped. Geometrically, the determinant represents the inverse of the density of lattice points in space (e.g., the number of lattice points in a large and sufficiently regular region of space A should be approximately equal to the volume of A divided by the determinant.) In particular, the determinant of a lattice does not depent on the choice of the basis. We will prove this formally later in this lecture.

The next simple upper bound on the determinant *(Hadamard inequality)* immediately follows from the fact that $\|\mathbf{b}_i^*\| \leq \|\mathbf{b}_i\|$.

Proposition 1. *For any lattice* $\mathcal{L}(\mathbf{B})$, $\det(\mathcal{L}(\mathbf{B})) \leq \prod \|\mathbf{b}_i\|$.

In the next lecture we will prove that the Gram-Schmidt orthogonalization of a basis can be computed in polynomial time. So, the determinant of a lattice can be computed in polynomial time by first computing the orthogonalized vectors $\mathbf{B}^*$, and then taking the product of their lengths. But there are simpler ways to express the determinant of a lattice that do not involve the Gram-Schmidt orthogonalized basis. The following proposition shows that the determinant of a lattice can be obtained from a simple matrix determinant computation. [2]

Proposition 2. *For any lattice basis* $\mathbf{B} \in \mathbb{R}^{n \times m}$

$$\det(\mathcal{L}(\mathbf{B})) = \sqrt{\det(\mathbf{B}^\top \mathbf{B})}.$$

In particular, if $\mathbf{B} \in \mathbb{R}^{n \times n}$ *is a (non-singular) square matrix then* $\det(\mathcal{L}(\mathbf{B})) = |\det(\mathbf{B})|$.

Proof. Remember the Gram-Schmidt orthogonalization procedure. In matrix notation, it shows that the orhogonalized vectors $\mathbf{B}^*$ satisfy $\mathbf{B} = \mathbf{B}^*\mathbf{T}$, where $\mathbf{T}$ is an upper triangular matrix with 1's on the diagonal, and the $\mu_{i,j}$ coefficients at position (j, i) for all $j < i$. So, our formula for the determinant of a lattice can be written as

$$\sqrt{\det(\mathbf{B}^\top \mathbf{B})} = \sqrt{\det(\mathbf{T}^\top \mathbf{B}^{*\top} \mathbf{B}^* \mathbf{T})} = \sqrt{\det(\mathbf{T}^\top) \det(\mathbf{B}^{*\top} \mathbf{B}^*) \det(\mathbf{T})}.$$

The matrices $\mathbf{T}^\top, \mathbf{T}$ are triangular, and their determinant can be easily computed as the product of the diagonal elements, which is 1. Now consider $\mathbf{B}^{*\top}\mathbf{B}^*$. This matrix is diagonal because the columns of $\mathbf{B}^*$ are orthogonal. So, its determinant can also be computed as the product of the diagonal elements which is

$$\det(\mathbf{B}^{*\top}\mathbf{B}^*) = \prod_i \langle \mathbf{b}_i^*, \mathbf{b}_i^* \rangle = (\prod_i \|\mathbf{b}_i^*\|)^2 = \det(\mathcal{L}(\mathbf{B}))^2.$$

Taking the square root we get $\sqrt{\det(\mathbf{T}^\top) \det(\mathbf{B}^{*\top} \mathbf{B}^*) \det(\mathbf{T})} = \det(\mathcal{L}(\mathbf{B}))$. $\square$

[2] Recall that the determinant of a matrix can be computed in polynomial time by computing $\det(\mathbf{B})$ modulo many small primes, and combining the results using the Chinese reminder theorem.

Now it is easy to show that the determinant does not depend on the particular choice of the basis, i.e., if two bases generate the same lattice then their lattice determinants have the same value.

Theorem 2. *Suppose* $\mathbf{B}$, $\mathbf{C}$ *are bases of the same lattice* $\mathcal{L}(\mathbf{B}) = \mathcal{L}(\mathbf{C})$. *Then,* $\det(\mathbf{B}) = \pm\det(\mathbf{C})$.

Proof. Suppose $\mathcal{L}(\mathbf{B}) = \mathcal{L}(\mathbf{C})$. Then $\mathbf{B} = \mathbf{C}\cdot\mathbf{U}$ where $\mathbf{U} \in \mathbb{Z}^{n\times n}$ is a unimodular matrix. Then $det(\mathbf{B}^\top\mathbf{B}) = \det((\mathbf{CU})^\top(\mathbf{CU})) = \det(\mathbf{U}^\top)\det(\mathbf{C}^\top\mathbf{B})\det(\mathbf{U}) = \det((\mathbf{C}^\top\mathbf{B})$ because $\det(\mathbf{U}) = 1$. □

We conclude this section showing that although not every lattice has an orthogonal basis, every integer lattice contains an orthogonal sublattice.

Theorem 3. *For any nonsingular* $B \in \mathbb{Z}^{n\times n}$, *let* $d = |\det(\mathbf{B})|$. *Then* $d\cdot\mathbb{Z}^n \subseteq \mathcal{L}(\mathbf{B})$.

Proof. Let $\mathbf{v}$ be any vector in $d\cdot\mathbb{Z}^n$. We know $\mathbf{v} = d\cdot\mathbf{y}$ for some integer vector $\mathbf{y}\in\mathbb{Z}^n$. We want to prove that $\mathbf{v}\in\mathcal{L}(\mathbf{B})$, i.e., $d\cdot\mathbf{y} = \mathbf{B}\cdot\mathbf{x}$ for some integer vector $\mathbf{x}$. Since $\mathbf{B}$ is non-singular, we can always find a solution $\mathbf{x}$ to the system $\mathbf{B}\cdot\mathbf{x} = d\cdot\mathbf{y}$ over the reals. We would like to show that $\mathbf{x}$ is in fact an integer vector, so that $d\mathbf{y}\in\mathcal{L}(\mathbf{B})$. We consider the elements x_i and use Cramer's rule:

$$\begin{aligned} x_i &= \frac{\det\left([\mathbf{b_1}, ..., \mathbf{b_{i-1}}, d\mathbf{y}, \mathbf{b_{i+1}}, ..., \mathbf{b_n}]\right)}{\det(B)} \\ &= \frac{d\cdot\det\left([\mathbf{b_1}, ..., \mathbf{b_{i-1}}, \mathbf{y}, \mathbf{b_{i+1}}, ..., \mathbf{b_n}]\right)}{\det(B)} \\ &= \det\left([\mathbf{b_1}, ..., \mathbf{b_{i-1}}, \mathbf{y}, \mathbf{b_{i+1}}, ..., \mathbf{b_n}]\right) \in \mathbb{Z} \end{aligned}$$

So, $\mathbf{x}$ is an integer vector. □

We may say that any integer lattice $\mathcal{L}(\mathbf{B})$ is periodic modulo the determinant of the lattice, in the sense that for any two vectors $\mathbf{x},\mathbf{y}$, if $\mathbf{x}\equiv\mathbf{y} \pmod{\det(\mathcal{L}(\mathbf{B}))}$, then $\mathbf{x}\in\mathcal{L}(\mathbf{B})$ if and only if $\mathbf{y}\in\mathcal{L}(\mathbf{B})$.

5. Minimum Distance

Definition 6. For any lattice $\Lambda = \mathcal{L}(\mathbf{B})$, the minimum distance of Λ is the smallest distance between any two lattice points:

$$\lambda(\Lambda) = \inf\{\|\mathbf{x}-\mathbf{y}\| : \mathbf{x},\mathbf{y}\in\Lambda, \mathbf{x}\neq\mathbf{y}\}.$$

We observe that the minimum distance can be equivalently defined as the length of the shortest nonzero lattice vector:

$$\lambda(\Lambda) = \inf\{\|\mathbf{v}\| : \mathbf{v}\in\Lambda\setminus\{\mathbf{0}\}\}.$$

This follows from the fact that lattices are additive subgroups of $\mathbb{R}^n$, i.e., they are closed under addition and subtraction. So, if $\mathbf{x}$ and $\mathbf{y}$ are distinct lattice points, then $\mathbf{x}-\mathbf{y}$ is a nonzero lattice point.

The first thing we want to prove about the minimum distance is that it is always achieved by some lattice vector, i.e., there is a lattice vector $\mathbf{x}\in\Lambda$ of length exactly $\|\mathbf{x}\| = \lambda(\Lambda)$. To prove this, we need first to establish a lower bound on $\lambda(\Lambda)$.

Theorem 4. *For every lattice basis* $\mathbf{B}$ *and its Gram-Schmidt orthogonalization* $\mathbf{B}^*$, $\lambda(\mathcal{L}(\mathbf{B})) \geq min\|\mathbf{b}_i^*\|$.

Proof. Note that $\mathbf{b}_i^*$ are not lattice vectors. Let us consider a generic lattice vector

$$\mathbf{B}\mathbf{x} \in \mathcal{L}(\mathbf{B}) \setminus \{\mathbf{0}\},$$

where $\mathbf{x} \in \mathbb{Z}^n \setminus \{\mathbf{0}\}$ and let k be the biggest index such that $x_k \neq 0$. We prove that

$$\|\mathbf{B}\mathbf{x}\| \geq \|\mathbf{b}_k^*\| \geq \min_i \|\mathbf{b}_i^*\|. \tag{1}$$

In order to prove (1), we take the scalar product of our lattice vector and $\mathbf{b}_k^*$. Using the orthogonality of $\mathbf{b}_k^*$ and $\mathbf{b}_i$ (for $i < k$) we get

$$\langle \mathbf{B}\mathbf{x}, \mathbf{b}_k^* \rangle = \sum_{i \leq k} \langle \mathbf{b}_i x_i, \mathbf{b}_k^* \rangle = x_k \langle \mathbf{b}_k, \mathbf{b}_k^* \rangle = x_k \|\mathbf{b}_k^*\|^2.$$

By Cauchy-Shwartz,

$$\|\mathbf{B}\mathbf{x}\| \cdot \|\mathbf{b}_k^*\| \geq |\langle \mathbf{B}\mathbf{x}, \mathbf{b}_k^* \rangle| \geq |x_k| \cdot \|\mathbf{b}_k^*\|^2.$$

Using $|x_k| \geq 1$ and dividing by $\|\mathbf{b}_k\|^*$, we get $\|\mathbf{B}\mathbf{x}\| \geq \|\mathbf{b}_k^*\|$. □

An immediate consequence of Theorem 4 is that for any matrix $\mathbf{B}$ with full column rank, the set $\mathcal{L}(\mathbf{B})$ is a lattice according to definition Definition

Notice that the lower bound $\min_i \|\mathbf{b}_i^*\|$ depends on the choice of the basis. We will see later in the course that some bases give better lower bounds than others, but at this point any nonzero lower bound will suffice. We want to show that there is a lattice vector of length λ. Consider a sphere of radius 2λ. Clearly, in the definition of $\lambda = \inf\{\|\mathbf{x}\| : \mathbf{x} \in \Lambda \setminus \{0\}\}$, we can restrict $\mathbf{x}$ to range over all lattice vectors inside this sphere. We observe that (by a volume argument) the sphere contains only finitely many lattice points. (Details below.) It follows that we can replace the inf operation with a min, and there is a point in the set achieving the smallest possible norm.

How can we use a volume argument, when points have volume 0? Put an open sphere of radius $\lambda/2$ around each lattice point. Since lattice points are at distance at least λ, the spheres are disjoint. The spheres with centers in S are also contained in a sphere S' of radius 3λ. So, since the volume of the small spheres (which is proportional to $1/2^n$) cannot exceed the volume of the big sphere S' (which has volume proportional to 3^n), there are at most 6^n lattice points.

6. Minkowski's theorem

We now turn to estimating the value of λ from above. Clearly, for any basis $\mathbf{B}$, we have $\lambda(\mathbf{B}) \leq \min_i \|\mathbf{b}_i\|$, because each column of $\mathbf{B}$ is a nonzero lattice vector. We would like to get a better bound, and, specifically, a bound that does not depend on the choice of the basis. We will prove an upper bound of the form $\lambda(\Lambda) \leq \alpha(n) \det(\Lambda)^{1/n}$.

Why $\det(\Lambda)^{1/n}$? The reason we look for bounds of this form is that the expression $\det(\Lambda)^{1/n}$ scales linearly with the lattice, i.e., if we multiply a lattice by a factor c, then we obtain $\lambda(c\Lambda) = c\lambda(\Lambda)$ and $\det(c\Lambda)^{1/n} = c\det(\Lambda)^{1/n}$.

The upper bound on $\lambda(\Lambda)$ we are going to prove was originally proved by Minkowski. Here we follow a different approach, by first proving a theorem of Blichfeldt from which Minkowski's theorem can be easily derived as a corollary.

Theorem 5. *Given a lattice* $\mathcal{L}(B)$ *and a set* $S \subseteq \mathbb{R}^m$ *if vol*$(S) > \det(B)$ *then* S *contains two points* $z_1, z_2 \in S$ *such that* $\mathbf{z}_1 - \mathbf{z}_2 \in \mathcal{L}(B)$.

Proof. Consider the sets $S_\mathbf{x} = S \cap (\mathbf{x} + \mathcal{P}(B))$, where $\mathbf{x} \in \mathcal{L}(B)$. Notice that these sets form a partition of S, i.e., they are pairwise disjoint and

$$S = \bigcup_{x \in \mathcal{L}(B)} S_x.$$

In particular we have

$$\mathrm{vol}(S) = \sum_{\mathbf{x} \in \mathcal{L}(B)} \mathrm{vol}(S_\mathbf{x}).$$

Notice that the transtated sets $S_\mathbf{x} - \mathbf{x} = (S - \mathbf{x}) \cap \mathcal{P}(B)$ are all contained in $\mathcal{P}(B)$. We want to prove that the $S_\mathbf{x}$ cannot be all mutually disjoint. Since $\mathrm{vol}(S_x) = \mathrm{vol}(S_\mathbf{x} - \mathbf{x})$, we have

$$\mathrm{vol}(\mathcal{P}(B)) < \mathrm{vol}(S) = \sum_{x \in \mathcal{L}B} \mathrm{vol}(S_x) = \sum_{x \in \mathcal{L}B} \mathrm{vol}(S_\mathbf{x} - \mathbf{x}).$$

The facts that $S_\mathbf{x} - \mathbf{x} \subseteq \mathcal{P}(B)$ and $\sum_{x \in \mathcal{L}(B)} \mathrm{vol}(S_\mathbf{x} - \mathbf{x}) > \mathrm{vol}(\mathcal{P}(B))$ imply that these sets cannot be disjoint, i.e. there exist two distinct vectors $\mathbf{x} \neq \mathbf{y} \in \mathcal{L}(B)$ such that $(S_\mathbf{x} - \mathbf{x}) \cap (S_\mathbf{y} - \mathbf{y}) \neq 0$.

Let $\mathbf{z}$ be any vector in the (non-empty) intersection $(S_\mathbf{x} - \mathbf{x}) \cap (S_\mathbf{y} - \mathbf{y})$ and define

$$\mathbf{z}_1 = \mathbf{z} + \mathbf{x} \in S_\mathbf{x} \subseteq S$$

$$\mathbf{z}_2 = \mathbf{z} + \mathbf{y} \in S_\mathbf{y} \subseteq S.$$

These two vectors satisfy

$$\mathbf{z}_1 - \mathbf{z}_2 = \mathbf{x} - \mathbf{y} \in \mathcal{L}(B).$$

□

As a corollary to Blichfeldt theorem we immediately get a result originally due to Minkowski that gives a bound on the length of the shortest vector in a lattice.

Corollary 1. *[Minkowski's convex body theorem] If* S *is a convex symmetric body of volume vol*$(S) > 2^m \det(B)$, *then* S *contains a non-zero lattice point.*

Proof. Consider the set $S/2 = \{\mathbf{x} : 2\mathbf{x} \in S\}$. The volume of $S/2$ satisfies

$$\mathrm{vol}(S/2) = 2^{-m}\mathrm{vol}(S) > \det(B)$$

By Blichfeldt theorem there exist $z_1, z_2 \in S/2$ such that $\mathbf{z}_1 - \mathbf{z}_2 \in \mathcal{L}(B) \setminus \{\mathbf{0}\}$. By definition of S/2, $2\mathbf{z}_1, 2\mathbf{z}_2 \in S$. Since S is symmetric, also $-2\mathbf{z}_2 \in S$ and by convexity,

$$\mathbf{z}_1 - \mathbf{z}_2 = \frac{2\mathbf{z}_1 - 2\mathbf{z}_2}{2} \in S$$

is a non-zero lattice vector contained in the set S. □

The relation between Minkowski theorem and bounding the length of the shortest vector in a lattice is easily explained. Consider first the ℓ_∞ norm: $\|\mathbf{x}\| = \max_i |x_i|$. We show that every (full rank, n-dimensional) lattice Λ always contains a nonzero vector $\|\mathbf{x}\| \leq \det(\Lambda)^{1/n}$. Let $l = \min\{\|\mathbf{x}\|_\infty : \mathbf{x} \in \Lambda \setminus \{\mathbf{0}\}\}$ and assume for contradition $l > \det(\Lambda)^{1/n}$. Take the hypercube $C = \{\mathbf{x} : \|\mathbf{x}\| < l\}$. Notice that C is convex, symmetric, and has volume $\mathrm{vol}(C) = (2l)^n >$

$2^n \det(\Lambda)$. So, by Minkowski's theorem, C contains a nonzero lattice vector $\mathbf{x}$. By definition of C, we have $\|\mathbf{x}\|_\infty < l$, a contradiction to the minimality of l.

For any full dimensionsional $\mathcal{L}(B)$ there exists a lattice point $x \in \mathcal{L}(B)/\{0\}$ such that

$$\|\mathbf{x}\|_\infty \leq \det(\mathbf{B})^{1/n}.$$

Using inequality $\|\mathbf{x}\| \leq \sqrt{n}\|\mathbf{x}\|_\infty$ (valid for any n-dimensional vector $\mathbf{x}$), we get a corresponding bound in the ℓ_2 norm. It is easy to see that for Euclidean norm the full dimensionality condition is not necessary because one can project a lattice of rank k to $\mathbb{R}^k$ while preserving the minimum distance.

Corollary 2. *For any lattice $\mathcal{L}(\mathbf{B})$ there exists a lattice point $x \in \mathcal{L}(\mathbf{B}) \setminus \{0\}$ such that*

$$\|\mathbf{x}\|_2 \leq \sqrt{n} \det(\mathbf{B})^{1/n}.$$

We could have proved the bound for the Euclidean norm directly, using a sphere instead of a cube, and then plugging in the formula for the volume of an n-dimensional sphere. This can be useful to get slighly better bounds. For example, in two dimensions, for any lattice Λ, the disk $S = \{\mathbf{x}\colon \|\mathbf{x}\| < \lambda(\Lambda)\}$ contains no nonzero lattice point. So, by Minkowki's theorem, the area of S can be at most $2^n \det(\Lambda) = 4 \det(\Lambda)$. But we know that the area of S is $\pi\lambda^2$. So, $\lambda(\Lambda) \leq 2\sqrt{\det(\Lambda)/\pi}$, which is strictly smaller than $\sqrt{2}\det(\Lambda)^{1/n}$.

We remark that a lattice Λ can contain vectors arbitrarily shorter than Minkowski's bound $\sqrt{n}\det(\Lambda)^{1/n}$. Consider for example the two dimensional lattice generated by the vectors $(1,0)^T$ and $(0,N)^T$, where N is a large integer. The lattice contains a short vector of length $\lambda = 1$. However, the determinant of the lattice is N, and Minkowski's bound $\sqrt{2}N^{1/2}$ is much larger than 1.

It can also be shown that Minkowski's bound cannot be asymptotically improved, in the sense that there is a constant c such that for any dimension n there is a n-dimensional lattice Λ_n such that $\lambda(\Lambda) > c\sqrt{n}\det(\Lambda)^{1/n}$. (See homework assignment.) So, up to constant factors, $O(\sqrt{n})\det(\Lambda)^{1/n}$ is the best upper bound one can possibly prove on the length of the shortest vector of any n-dimensional lattice.

7. A simple application

As an application of Minkowski's theorem we show that any prime number p congruent to 1 mod 4 can be written as the sum of two squares.

Theorem 6. *For every prime $p \equiv 1 \bmod 4$ there exist integers $a, b \in \mathbb{Z}$ such that $p = a^2 + b^2$*

Proof. Let $p \in \mathbb{Z}$ be a prime such that $p \equiv 1 \bmod 4$. Then $\mathbb{Z}_p^*$ is a group such that $4 \mid o(\mathbb{Z}_p^*) = p - 1$. Therefore, there exists an element of multiplicative order 4, and -1 is a quadratic residue modulo p, i.e. there exists an integer i such that $i^2 \equiv -1 \pmod{p}$. It immediately follows that

$$p \mid i^2 + 1. \tag{2}$$

Now define the lattice basis

$$\mathbf{B} = \begin{bmatrix} 1 & 0 \\ i & p \end{bmatrix}.$$

By Minkowski's theorem there exists an integer vector $\mathbf{x}$ such that $0 < \|\mathbf{Bx}\|_2 < \sqrt{2} \cdot \sqrt{\det(\mathbf{B})}$. Squaring this equation yields

$$0 < \|\mathbf{Bx}\|_2^2 < 2 \cdot \det(\mathbf{B}) = 2p. \tag{3}$$

The middle term expands to

$$\left\| \begin{bmatrix} x_1 \\ ix_1 + px_2 \end{bmatrix} \right\|^2 = x_1^2 + (ix_1 + px_2)^2 \tag{4}$$

If we let $a = x_1$ and $b = ix_1 + px_2$, (2) becomes

$$0 < a^2 + b^2 < 2p \tag{5}$$

Hence if we can show that $a^2 + b^2 \equiv 0 \bmod p$, by necessity $a^2 + b^2 = p$. Expanding the right side of (3) produces $x_1^2 + i^2x_1^2 + p^2x_2^2 + 2ix_1px_2$, which can be factored into

$$p(px_2^2 + 2ix_1x_2) + x_1^2(i^2 + 1)$$

Obviously p divides the first term, and by (1) p divides the second term. Thus $a^2 + b^2 \equiv 0 \bmod p$, and therefore by (4) $a^2 + b^2 = p$. □

This application shows how lattices can be used to prove non-trivial facts in number theory. A similar theorem that can be proved with the same lattice techniques is the following.

Theorem 7. $\forall n \in \mathbb{Z}^+\ \exists a, b, c, d \in \mathcal{Z} : n = a^2 + b^2 + c^2 + d^2$.

The proof is left to the reader as an excercise. As you can easily guess, the proofs involves a 4-dimensional lattice.

8. Successive Minima

Definition 7. For any lattice Λ and integer $k \leq \operatorname{rank}(\Lambda)$, let $\lambda_k(\Lambda)$ be the smallest $r > 0$ such that Λ contains at least k linearly independent vectors of length boounded by r.

The successive minima of a lattice generalize the minimum distance $\lambda = \lambda_1$. By the same volume argument used to show that there exists vectors of length λ, one can show that there exist (linearly independent) lattice vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ of lengths $\lambda_1, \ldots, \lambda_k$. Minkowski's theorem can also be generalized to provide a bound not just on λ_1, but on the geometric mean of all successive minima.

Theorem 8. *For any lattice L, $\prod_{i=1}^{n} \lambda_i \leq \frac{2^n \det(L)}{vol(S_n)}$, where S_n is the n-dimensional unit ball.*

Proof. Assume for contradiction this is not the case, i.e., $\prod_i \lambda_i > 2^n \det(L)/\mathrm{vol}(S_n)$ and let $\mathbf{x_1}, \ldots, \mathbf{x_n}$ be linearly independent vectors such that $\|\mathbf{x_i}\| = \lambda_i$. Consider the orthogonalized vectors $\mathbf{x_i}^*$ and define the transformation

$$T(\sum c_i \mathbf{x_i}^*) = \sum \lambda_i c_i \mathbf{x_i}^*$$

that expands coordinate $\mathbf{x_i}^*$ by the factor λ_i. If we apply T to the open unit ball S_n we get a symmetric convex body $T(S_n)$ of volume $(\prod_i \lambda_i)\mathrm{vol}(S_n) > 2^n \det(L)$. By Minkowski's first theorem $T(S_n)$ contains a lattice point $\mathbf{y} = T(\mathbf{x})$ (with $\|\mathbf{x}\| < 1$) different from the origin. Let $\mathbf{x} = \sum c_i \mathbf{x_i}^*$ and $\mathbf{y} = \sum \lambda_i c_i \mathbf{x_i}^*$. Since $\mathbf{y}$ is not zero, some c_i is not zero. Let k the largest index such that $c_i \neq 0$. Notice that $\mathbf{y}$ is linearly independent from $\mathbf{x_1}, \ldots, \mathbf{x_{k-1}}$

because $\langle \mathbf{x_k}^*, \mathbf{y} \rangle = \lambda_k c_k \|\mathbf{x_k}^*\|^2 > 0$. We now show that $\|\mathbf{y}\| < \lambda_k$, contradicting the definition of λ_i for some $i \leq k$.

$$\|\mathbf{y}\|^2 = \sum_{i \leq k} \lambda_i^2 c_i^2 \|\mathbf{x_i}^*\|^2 \leq \sum_{i \leq k} \lambda_k^2 c_i^2 \|\mathbf{x_i}^*\|^2 = \lambda_k^2 \|\mathbf{x}\|^2 < \lambda_k^2$$

□