Lattices

Some definitions

Throughout this section k and n will be positive integers with $k \leq n$.

A **lattice** L, of **rank** (or **dimension**) k in $\mathbb{R}^n$ is a $\mathbb{Z}$-module spanned by k linearly independent vectors in $\mathbb{R}^n$ (linearly independent over $\mathbb{R}$). Thus

$$L = \{\sum_{i=1}^{k} x_i b_i : x_i \in \mathbb{Z}\},$$

where $b_1, \ldots, b_k$ are linearly independent column vectors in $\mathbb{R}^n$. Any such linearly independent spanning set for L is called a **basis**.

If $c_1, \ldots, c_k$ is another basis, then there exists a $k \times k$ integer matrix U, with determinant ± 1, such that

$$(c_1, \ldots, c_k) = (b_1, \ldots, b_k)U, \quad (*)$$

where $(c_1, \ldots, c_k)$ and$(b_1, \ldots, b_k)$ are $n \times k$ matrices.

The determinant of L, denoted $\Delta(L)$ or $\det(L)$, is defined by

$$\Delta(L) = \det(B^t B)^{1/2}$$

where $B = (b_1, \ldots, b_k)$. The matrix $B^t B$ will be a positive-definite symmetric matrix, and hence will have positive determinant. It takes the form

$$\left(b_i^t b_j\right)_{i,j \leq k}$$

with entries being the scalar products of the basis vectors. One sees from (*) that the definition of $\Delta(L)$ is independent of the choice of basis. Moreover in the special case $k = n$ we have

$$\Delta(L) = |\det(b_1, \ldots, b_n)| = |\det(B)|.$$

In fact $\Delta(L)$ is the k-dimensional volume of the **fundamental parallelotope**

$$\{\sum_{i=1}^{k} x_i b_i : 0 < x_i \leq 1\}.$$

If $L_1 \subseteq L_2$ we say that L_1 is a **sublattice** of L_2. (Here the corresponding values of n must be the same, but we might have $k_1 < k_2$.)

Example Let $k = n = 2$ and take

$$b_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad b_2 = \begin{pmatrix} 2 \\ 0 \end{pmatrix}.$$

Then

$$\Delta(L) = \left| \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} \right| = 2.$$

The symmetric matrix $Q = B^t B$ may be used to define a positive definite quadratic form

$$Q(x, y) = (x\ y)B^t B \begin{pmatrix} x \\ y \end{pmatrix} = (x\ y) \begin{pmatrix} 2 & 2 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 2x^2 + 4xy + 4y^2.$$

In general any positive definite quadratic form arises from a suitable lattice in this way. Note that $Q(x_1, \ldots, x_k) = || \sum_{i=1}^{k} x_i b_i ||^2$, where $||\ ||$ is the standard Euclidean norm on $\mathbb{R}^n$.

Instead of the basis b_1, b_2 we could use

$$c_1 = \begin{pmatrix} 7 \\ 19 \end{pmatrix}, \quad c_2 = \begin{pmatrix} 6 \\ 16 \end{pmatrix},$$

which is also a basis for L. However in many situations we want to find a basis in which the vectors are as short as possible.

The shortest non-zero lattice vector

A lattice L has a well-defined minimal length among non-zero vectors. The *square* of this length is usually denoted by $M_1(L)$. To prove this assertion we consider the quadratic form Q associated to L. Since Q is positive definite one may complete the square successively to produce

$$\begin{aligned} Q(x_1, \ldots, x_k) = a_1(x_1 + \lambda_{12}x_2 + \ldots + \lambda_{1k}x_k)^2 \\ + a_2(x_2 + \lambda_{23}x_3 + \ldots + \lambda_{2k}x_k)^2 + \ldots + a_k x_k^2. \end{aligned}$$

The coefficients a_j will all be strictly positive since Q is positive definite. Then the vector b_1 has $||b_1||^2 = Q(1, 0, \ldots, 0) = a_1$. We claim that there are only finitely many lattice vectors b with $||b||^2 \leq a_1$, or in other words, only finitely many integer vectors $(x_1, \ldots, x_k)$ with $Q(x_1, \ldots, x_k) \leq a_1$. This condition is equivalent to

$$a_1(x_1 + \lambda_{12}x_2 + \ldots + \lambda_{1k}x_k)^2 + a_2(x_2 + \lambda_{23}x_3 + \ldots + \lambda_{2k}x_k)^2 + \ldots + a_k x_k^2 \leq a_1.$$

Thus we have $a_k x_k^2 \leq a_1$, which shows that there are finitely many possibilities for x_k. Then $a_{k-1}(x_{k-1} + \lambda_{k-1,k} x_k)^2 + a_k x_k^2 \leq a_1$, so that there are finitely many possibilities for x_{k-1}, and so on.

Example
Continuing the example before, we have

$$Q(x, y) = 2x^2 + 4xy + 4y^2 = 2(x + y)^2 + 2y^2,$$

with $||b_1||^2 = a_1 = 2$. If $2(x + y)^2 + 2y^2 \leq 2$ then $y^2 \leq 1$ so that $y = 0$ or ± 1. Trying these values of y we find that $Q(x, y) \leq 2$ for integers x, y precisely when $(x, y) = (0, 0), (1, 0), (-1, 0), (-1, 1)$ or $(1, -1)$. Since we want to have a non-zero vector the first of these is disallowed, leaving 4 vectors whose length is the minimal value $\sqrt{2}$, namely $b_1, -b_1, -b_1 + b_2$ and $b_1 - b_2$. In particular $M_1(L) = 2$.

In general, in higher dimensions, this process is horribly inefficient, particularly if one starts with a basis that consists of vectors that are much larger than necessary. Thus we shall describe, later in this section, a good algorithm (the "LLL" algorithm) for finding a tolerably good basis, which can be used as a starting point for the above process. Indeed in many cases the LLL basis is already good enough for the required application.

Hermite's theorem

Hermite's theorem states that for each $k \in \mathbb{N}$ there is a constant μ_k such that

$$M_1(L)^k \leq \mu_k \Delta(L)^2$$

for every lattice L of rank k.

Sketch proof
Since we may restrict attention to the k-dimensional subspace of $\mathbb{R}^n$ in which L lies it suffices to suppose that $k = n$. Consider the fundamental parallelepiped

$$P = \{\sum_{i=1}^{k} x_i b_i : 0 < x_i \leq 1\}.$$

For different lattice points x the translates $P + x$ are disjoint, and their union is the whole of $\mathbb{R}^k$. Thus in a large cube C the number of lattice points is approximately $\mathrm{Vol}(C)/\mathrm{Vol}(P) = \mathrm{Vol}(C)/\Delta(L)$. This approximation gets better and better the larger the cube we use.

On the other hand, consider the sphere

$$S := \{x \in \mathbb{R}^k : ||x|| < \frac{1}{2}\sqrt{M_1(L)}\}.$$

We claim that the translates $S+x$ of these are also disjoint, for distinct lattice vectors $x \in L$. If this were not the case we would have two different vectors x, x' a distance less than $\sqrt{M_1(L)}$ apart. This would produce a non-zero vector $x - x' \in L$ of length less than $\sqrt{M_1(L)}$, which is impossible.

Since the translates of S are disjoint, the number of them which can lie in a large cube C is at most $\mathrm{Vol}(C)/\mathrm{Vol}(S)$. However, since there is (essentially) one such translate for every lattice point in C we deduce that $\mathrm{Vol}(C)/\Delta(L)$ is (more or less) at most $\mathrm{Vol}(C)/\mathrm{Vol}(S)$. Again the approximation gets better and better the larger the cube we use. We deduce that $\Delta(L)^{-1} \leq \mathrm{Vol}(S)^{-1}$. If we write $\mu_k^{-1/2}$ for the volume of the sphere of radius $1/2$ we have $\mathrm{Vol}(S) = M_1(L)^{k/2}\mu_k^{-1/2}$, and the result follows.

This argument does not give the best possible value for μ_k. Finding the optimal values is a well-known problem, which has been solved only for $k \leq 8$.

An important corollary

Corollary For any lattice L there is a real number $c(L) > 0$ such that $\Delta(L_1) \geq c(L)$ for every sublattice L_1 of L, irrespective of the rank of L_1.

Proof Suppose that L_1 is a sublattice of L and that L_1 has rank h. Then $M_1(L) \leq M_1(L_1)$, and $M_1(L_1) \leq \mu_h^{1/h}\Delta(L_1)^{2/h}$. Thus

$$\Delta(L_1) \geq M_1(L)^{h/2}\mu_h^{-1/2},$$

so that if L has rank k we can take $c(L) = \min_{h \leq k} M_1(L)^{h/2}\mu_h^{-1/2}$.

The Gram–Schmidt process

The Gram–Schmidt process deals with vector spaces over $\mathbb{R}$, and converts an arbitrary basis into an orthogonal one. In our situation we do not need to produce an orthonormal basis (with vectors of length 1) but merely one in which any two different basis vectors are orthogonal. To fix our notation let us recall how this is done.

Given linearly independent vectors $b_1, \ldots, b_k$ in $\mathbb{R}^n$, one defines $b_1^* = b_1$ and

$$b_i^* = b_i - \sum_{j=1}^{i-1} \mu_{ij} b_j^* \quad (2 \leq i \leq k)$$

with

$$\mu_{ij} = \frac{\langle b_i, b_j^* \rangle}{\langle b_j^*, b_j^* \rangle} \quad (1 \leq j \leq i-1).$$

Then $b_1^*, \ldots, b_k^*$ will be orthogonal, and will span the same vector space as $b_1, \ldots, b_k$. However they will not span the same *lattices*, unless by fluke all the numbers μ_{ij} are integers.

LLL-reduced bases

We now begin the study of the reduction process discovered by Lenstra, Lenstra and Lovász. A lattice basis $b_1, \ldots, b_k$ is said to be LLL-reduced if the associated Gram–Schmidt basis and associated coefficients μ_{ij} satisfy

(1) $|\mu_{ij}| \leq \frac{1}{2}$ for $1 \leq j < i \leq k$; and

(2) $||b_i^* + \mu_{i,i-1} b_{i-1}^*||^2 \geq \frac{3}{4} ||b_{i-1}^*||^2$ for $2 \leq i \leq k$.

Motivation

Condition (1) says that $b_1, \ldots, b_k$ are "almost orthogonal", in the sense that the Gram–Schmidt coefficients μ_{ij} are small.

Condition (2) restricts the relative sizes of the basis vectors. Using orthogonality one can see that (2) is equivalent to

$$||b_i^*||^2 \geq (\tfrac{3}{4} - \mu_{i,i-1}^2) ||b_{i-1}^*||^2 \quad (2 \leq i \leq k).$$

Thus the lengths $||b_i^*||$ form an "almost-increasing" sequence.

The precise fraction $\frac{3}{4}$ which appears in (2) is not crucial. Any value strictly between $\frac{1}{4}$ and 1 could be used, but $\frac{3}{4}$ is a good compromise. Larger values would give better bases, but at the expense of longer running times.

Properties of LLL-reduced bases

Throughout this section suppose that $b_1, \ldots, b_k$ is an LLL-reduced basis for L, with associated Gram–Schmidt coefficients μ_{ij}. Then

(a) We have
$$||b_j||^2 \leq 2^{i-1}||b_i^*||^2 \leq 2^{i-1}||b_i||^2$$
for $1 \leq j \leq i \leq k$.

(b) $\Delta(L) \leq \prod_{i=1}^{k} ||b_i|| \leq 2^{k(k-1)/4}\Delta(L)$.

(c) $||b_1|| \leq 2^{(k-1)/4}\Delta(L)^{1/k}$.

(d) For every non-zero vector x in L we have $||b_1|| \leq c||x||$ where
$$c = \max_{1 \leq i \leq k} \frac{||b_1||}{||b_i^*||} \leq 2^{(k-1)/2}.$$

Remarks

From (c) one sees that
$$M_1(L)^k \leq ||b_1||^{2k} \leq 2^{k(k-1)/2}\Delta(L)^2,$$
giving us a form of Hermite's theorem, with the (rather bad) explicit value $\mu_k = 2^{k(k-1)/2}$. Notice also that (d) shows that $||b_1||$ is close to being a minimal-length vector in the lattice. It is larger than minimal by at worst the constant factor c.

Proofs

(a) We have
$$||b_i^*||^2 \geq (\tfrac{3}{4} - \mu_{i,i-1}^2)||b_{i-1}^*||^2 \geq \tfrac{1}{2}||b_{i-1}^*||^2 \quad (2 \leq i \leq k).$$
Hence $||b_j^*||^2 \leq 2^{i-j}||b_i^*||^2$ for $1 \leq j \leq i \leq k$ by induction. Thus
$$||b_i||^2 \;=\; ||b_i^*||^2 + \sum_{j=1}^{i-1} \mu_{ij}^2||b_j^*||^2 \quad (**)$$

$$\begin{aligned} &\leq \left(1 + \sum_{j=1}^{i-1} 2^{i-j-2}\right) ||b_i^*||^2 \\ &= (1 + \tfrac{1}{4}(2^i - 2))||b_i^*||^2 \\ &\leq 2^{i-1}||b_i^*||^2, \end{aligned}$$

so that

$$||b_j||^2 \leq 2^{j-1}||b_j^*||^2 \leq 2^{j-1+i-j}||b_i^*||^2 = 2^{i-1}||b_i^*||^2.$$

The inequality (a) then follows, using (**).

(b) Let B be the matrix with columns $b_1, \ldots, b_k$ and let B^* be the matrix with columns $b_1^*, \ldots, b_k^*$. Then $B = B^*J$ where J is the upper triangular matrix

$$J = \begin{pmatrix} 1 & \mu_{2,1} & \mu_{3,1} & \cdots & \mu_{k,1} \\ 0 & 1 & \mu_{2,3} & \cdots & \mu_{k,2} \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & \mu_{k,k-1} \\ 0 & \cdots & \cdots & \cdots & 1 \end{pmatrix}.$$

Then $\Delta(L)^2 = \det(B^tB) = \det((B^*)^tB^*)$, since $\det(J) = 1$. However the vectors b_i^* are orthogonal, so that $(B^*)^tB^*$ is a diagonal matrix, with diagonal entries $(b_i^*)b_i^* = ||b_i^*||^2$. We therefore conclude that $\Delta(L)^2 = \prod_{i=1}^k ||b_i^*||^2$. However (**) yields $||b_i^*|| \leq ||b_i||$, so that

$$\begin{aligned} \Delta(L) &\leq \prod_{i=1}^k ||b_i|| \quad \text{(which is the first inequality in (b))} \\ &\leq \prod_{i=1}^k 2^{(i-1)/2}||b_i^*|| \quad \text{(using (a) with } j = i) \\ &= 2^{k(k-1)/4} \prod_{i=1}^k ||b_i^*|| \\ &= 2^{k(k-1)/4}\Delta(L). \end{aligned}$$

Remark The same argument shows that $\prod_{i=1}^{k} ||c_i|| \geq \Delta(L)$ for any basis $c_1, \ldots, c_k$ of L, whether LLL-reduced or not.

(c) Set $j = 1$ in (a) and take the product for $1 \leq i \leq k$ to get

$$||b_1||^{2k} \leq \prod_{i=1}^{k} 2^{i-1}||b_i^*||^2 = 2^{k(k-1)/2}\Delta(L)^2,$$

and the required inequality follows.

(d) We have $||b_1||^2 \leq c^2||b_i^*||^2$ for $1 \leq i \leq k$, by definition of c. Any non-zero x in L may be written as $\sum_{i=1}^{k} r_i b_i$ with $r_i \in \mathbb{Z}$, and also as $\sum_{i=1}^{k} r_i^* b_i^*$ with $r_i^* \in \mathbb{R}$. Choose the largest index i such that $r_i \neq 0$ and call it $i = h$. By construction of the Gram–Schmidt basis one has $b_h^* - b_h \in \langle b_1, \ldots, b_{h-1} \rangle$, and also $b_i^* \in \langle b_1, \ldots, b_{h-1} \rangle$ for $1 \leq i \leq h-1$. It follows that $(r_h - r_h^*)b_h \in \langle b_1, \ldots, b_{h-1} \rangle$, and since the the vectors b_i are linearly independent this shows that $r_h^* = r_h$. However r_h is a non-zero integer so that $||r_h^*|| \geq 1$. We therefore have

$$||x||^2 = \sum_{i=1}^{k} (r_i^*)^2||b_i^*||^2 \geq (r_h^*)^2||b_h^*||^2 \geq ||b_h^*||^2 \geq ||b_1||^2/c^2$$

as required.