8.9 LLL algorithm

Although Dirichlet's theorems on simultaneous approximations are direct generalizations of Dirichlet's theorem for $n = 1$, when we speak about effective algorithms for finding rational numbers whose existence is guaranteed by those theorems, there is a significant difference between cases $n = 1$ and $n \geq 2$. Namely, for $n = 1$, as we have already seen, rational approximations with required property can be obtained by continued fractions. For $n \geq 2$, the situation is more complex. There are various generalizations of continued fractions and their properties to higher dimensions. We will focus on one, the most famous and for numerous applications the most important: the so-called LLL algorithm. This algorithm efficiently (in a polynomial time) provides rational approximations of a similar quality as the ones guaranteed by Dirichlet's theorems (some additional factors will appear compared to the optimal approximations provided by the theory). Additional information on the LLL algorithm and its application can be found in [279], [326], [378, Chapters 5 and 6] and [399].

The LLL algorithm is connected to the problem of finding the shortest non-zero vector in a lattice.

Definition 8.11. *Let n be a positive integer and let $b_1, \ldots, b_n$ be linearly independent vectors in $\mathbb{R}^n$. The* lattice *($\mathbb{Z}$-module) L determined by these vectors is the set of all their integer linear combinations*

$$L = \left\{ \sum_{i=1}^{n} x_i b_i \,:\, x_i \in \mathbb{Z} \right\}.$$

We call $B = \{b_1, \ldots, b_n\}$ a lattice basis *of L.*

For example, in $\mathbb{R}^2$, if $b_1 = (1, 0)$, $b_2 = (0, 1)$, then L is the lattice of all points in the plane with integer coordinates.

We will also denote by B the matrix whose columns are vectors $b_1, \ldots, b_n$. The determinant of the lattice L is defined by $\Delta(L) = |\det(B)|$. It is well-defined because lattice basis is unique up to multiplication on the right by matrices from $GL_n(\mathbb{Z})$, i.e. matrices with integer coefficients and determinant ± 1.

We denote by $\langle\, , \rangle$ the standard scalar product on $\mathbb{R}^n$. Note that the square of the Euclidean norm of a vector $v = \sum_{i=1}^{n} x_i b_i$ induces the quadratic form

$$\|v\|^2 = v^\tau v = x^\tau B^\tau B x = Q(x).$$

Since a lattice is a discrete set, the notion of the length of the shortest non-zero vector in a lattice is well-defined. From Minkowski's theorem on convex bodies, it follows that if C is a compact convex set, symmetric with respect to the origin, and if $\mu(C) \geq 2^n \Delta(L)$, then C contains a non-zero vector from L (apply Theorem 8.54 to the set $B^{-1}(C)$). If we take for C the n-dimensional ball, we obtain an upper bound for the length of the shortest non-zero vector in the lattice.

Proposition 8.56. *There is a constant γ_n such that*

$$\min_{v \in L \setminus \{0\}} \|v\| \leq \sqrt{\gamma_n} \Delta(L)^{1/n}.$$

Optimal values of γ_n are known for $n \leq 8$. For instance, $\gamma_1 = 1$, $\gamma_2 = \frac{4}{3}$, $\gamma_3 = 2$.

One lattice can have more bases, so we may ask, is it possible to choose a basis which will have some additional "good" properties? Clearly, B is a basis of the vector space $\mathbb{R}^n$. By the Gram-Schmidt process, we can obtain an orthogonal basis for the same vector space ($b_i^* = b_i - \sum_{j=1}^{i-1} \mu_{ij} b_j^*$, $i = 1, \ldots, n$, where $\mu_{ij} = \frac{\langle b_i, b_j^* \rangle}{\langle b_j^*, b_j^* \rangle}$). However, the new basis does not need to determine the same lattice as the initial basis B, because coefficients μ_{ij} need not be integers. In general, a lattice does not need to have an orthogonal basis. In 1982, A. K. Lenstra, H. W. Lenstra and L. Lovász introduced the notion of *LLL-reduced basis* with the following properties:

1) $|\mu_{i,j}| \leq \frac{1}{2}$, $1 \leq j < i \leq n$;

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

The first condition can be interpreted by saying that an LLL-reduced basis is "almost orthogonal", while the second condition states that the sequence of norms of vector $\|b_i^*\|$ is "almost increasing". An additional significant property of an LLL-reduced basis is that the first vector in that basis is very short, i.e. it has a small norm. It can be proved that $\|b_1\| \leq 2^{(n-1)/2} \|x\|$, for all non-zero vectors $x \in L$; however, in applications, it often occurs that $\|b_1\|$ is the shortest non-zero vector of L. We will provide more precise information in the following lemma.

Lemma 8.57. *Let $\{b_1, \ldots, b_n\}$ be an LLL-reduced basis, and $\{b_1^*, \ldots, b_n^*\}$ the corresponding Gram-Schmidt basis. Then*

1) $\|b_j\|^2 \leq 2^{i-1} \|b_i^*\|^2$, $1 \leq j \leq i \leq n$;

2) $\Delta(L) \le \prod_{i=1}^{n} \|b_i\| \le 2^{n(n-1)/4}\Delta(L)$*;*

3) $\|b_1\| \le 2^{(n-1)/4}\,(\Delta(L))^{1/n}$*;*

4) For every $x \in L$*,* $x \neq 0$*, we have* $\|b_1\|^2 \le c_1\|x\|^2$*, where*

$$c_1 = \max\left\{\frac{\|b_1\|^2}{\|b_i^*\|^2} \,:\, 1 \le i \le n\right\} \le 2^{n-1}.$$

5) For a vector $y \notin L$*, let us define* $\sigma = B^{-1}y$*, where* B *is the matrix whose columns are* $b_1, \ldots, b_n$*. Let* i_0 *be the largest index such that* $\sigma_{i_0} \notin \mathbb{Z}$*, and* $\|\sigma_{i_0}\|$ *the distance from* σ_{i_0} *to the nearest integer. Then for every* $x \in L$*, we have*

$$\|x - y\|^2 \ge c_1^{-1}\|\sigma_{i_0}\|^2\|b_1\|^2.$$

Proof:

1) From the definition of an LLL-reduced basis, it follows that

$$\|b_i^*\|^2 \ge (\frac{3}{4} - \mu_{i,i-1}^2)\|b_{i-1}^*\|^2 \ge \frac{1}{2}\|b_{i-1}^*\|^2$$

for $i = 1, 2, \ldots, n$. By induction, we decude that $\|b_j^*\|^2 \le 2^{i-j}\|b_i^*\|^2$ for $1 \le j \le i \le n$. Now, from the definition of Gram-Schmidt basis, we obtain

$$\|b_i\|^2 = \|b_i^*\|^2 + \sum_{j=1}^{i-1}\mu_{i,j}^2\|b_j^*\|^2 \le \Big(1 + \sum_{j=1}^{i-1}2^{i-j-2}\Big)\|b_i^*\|^2$$
$$= \Big(1 + \frac{1}{4}(2^i - 2)\Big)\|b_i^*\|^2 \le 2^{i-1}\|b_i^*\|^2.$$

Hence, for $1 \le j \le i \le n$,

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

2) From the orthogonality of the vectors b_i^*, it follows that

$$\Delta(L) = |\det(b_1^*, \ldots, b_n^*)| = \prod_{i=1}^{n}\|b_i^*\|.$$

From statement 1) and the inequality $\|b_i^*\| \le \|b_i\|$, we obtain

$$\Delta(L) \le \prod_{i=1}^{n}\|b_i\| \le \prod_{i=1}^{n}2^{(i-1)/2}\|b_i^*\| \le 2^{n(n-1)/4}\prod_{i=1}^{n}\|b_i^*\| = 2^{n(n-1)/4}\Delta(L).$$

3) Let us insert $j = 1$ into statement 1) and take the product over all possible i's. We obtain

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

and from this, the statement 3) follows.

4) Let

$$x = \sum_{i=1}^{n} r_i b_i = \sum_{i=1}^{n} r_i' b_i^*, \quad r_i \in \mathbb{Z},\ r_i' \in \mathbb{R}.$$

Let i_0 be the largest index for which $r_{i_0} \neq 0$. Since $b_1, \ldots, b_i$ determine the same vector space as $b_1^*, \ldots, b_i^*$, for all i, and since b_{i+1}^* is the projection of b_{i+1} to the orthogonal complement of that space, we conclude that $r_{i_0}' = r_{i_0}$. Therefore, we have

$$\|x\|^2 = \sum_{i=1}^{n} r_i'^2\|b_i^*\|^2 \geq {r_{i_0}'}^2\|b_{i_0}^*\|^2 \geq \|b_{i_0}^*\|^2 \geq c_1^{-1}\|b_1\|^2.$$

5) Let

$$x = \sum_{i=1}^{n} r_i b_i = \sum_{i=1}^{n} r_i' b_i^*, \quad y = \sum_{i=1}^{n} \sigma_i b_i = \sum_{i=1}^{n} \sigma_i' b_i^*.$$

If i_1 is the largest index such that $r_{i_1} \neq \sigma_{i_1}$, then $r_{i_1} - \sigma_{i_1} = r_{i_1}' - \sigma_{i_1}'$, so we have

$$\|x - y\|^2 \geq |r_{i_1} - \sigma_{i_1}|^2\|b_{i_1}^*\|^2 \geq |r_{i_1} - \sigma_{i_1}|^2 c_1^{-1}\|b_1\|^2.$$

If $i_1 < i_0$, then $\sigma_{i_0} = r_{i_0} \in \mathbb{Z}$, which is a contradiction. If $i_1 = i_0$, then we have $|r_{i_1} - \sigma_{i_1}| = |r_{i_0} - \sigma_{i_0}| \geq \|\sigma_{i_0}\|$ and we obtain the desired inequality. Finally, if $i_1 > i_0$, then $\sigma_{i_1} \in \mathbb{Z}$, so from $\sigma_{i_1} \neq r_{i_1}$ and $\|\sigma_{i_0}\| \leq \frac{1}{2}$, we obtain $|r_{i_1} - \sigma_{i_1}| \geq 1 \geq \|\sigma_{i_0}\|$. □

In the paper [271] from 1982, A. K. Lenstra, H. W. Lenstra and L. Lovász demonstrated a polynomial time algorithm for the construction of an LLL-reduced basis from an arbitrary lattice basis (named the *LLL algorithm* after them). The algorithm soon found numerous applications, e.g. in the factorization of polynomials with rational coefficients, the cryptanalysis of the RSA cryptosystem with a small public or private exponent, the knapsack-problem, Diophantine approximations and Diophantine equations.

In 1989, De Weger [416] proposed a variant of the LLL algorithm which uses only integer arithmetics (provided the input data are integers) and avoids problems with numerical stability of the algorithm.

In program package PARI, the LLL algorithm is implemented by the function `qflll`(x), which as a result returns the transformation matrix T such that xT is an LLL-reduced basis of the lattice generated by the columns of the matrix x. Functions for finding an LLL-reduced basis also exist in other program packages (e.g. `lattice` (Maple), `LatticeReduce` (Mathematica), `LLL` (Magma)).

Before we move on to applications of lattices to simultaneous Diophantine approximations in higher dimensions, let us consider if we can interpret the well-known problem of rational approximation of one irrational number in terms of lattices. We know that good rational approximations of an irrational number α can be obtained using convergents $\frac{p_i}{q_i}$ of the continued fraction of $\alpha = [a_0, a_1, a_2, \ldots]$. The convergents satisfy the recurrence

$$p_{i+1} = a_i p_i + p_{i-1}, \quad q_{i+1} = a_i q_i + q_{i-1}.$$

By the notation

$$M(i) = \begin{pmatrix} q_i & q_{i+1} \\ p_i & p_{i+1} \end{pmatrix}, \tag{8.44}$$

we can write the recurrence in terms of matrices as

$$M(i) = M(i-1) \cdot \begin{pmatrix} 0 & 1 \\ 1 & a_i \end{pmatrix}, \tag{8.45}$$

with $M(-1) = I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

The matrices $M(i)$ have determinant ± 1, i.e. $M(i) \in GL_2(\mathbb{Z})$. Let C be a positive real number. We denote by $L_C(\alpha)$ the lattice generated by the columns of the matrix $\begin{pmatrix} 1 & 0 \\ -C\alpha & C \end{pmatrix}$. We have

$$\begin{pmatrix} 1 & 0 \\ -C\alpha & C \end{pmatrix} M(i) = \begin{pmatrix} q_i & q_{i+1} \\ C(p_i - q_i\alpha) & C(p_{i+1} - q_{i+1}\alpha) \end{pmatrix}.$$

Let us recall that $|p_i - q_i\alpha| < \frac{1}{q_i}$. If we now choose $C \approx q_i^2$, then we see that the first column in the last matrix is the vector $(q_i, C(p_i - q_i\alpha))^\tau$ from the lattice $L_C(\alpha)$ which is significantly shorter than the vector of the initial lattice basis. Furthermore, q_i is the first component of that short vector.

Proposition 8.58. *Let α and $C > 0$ be real numbers. If $(u, C(w - \alpha u))^\tau$, $u > 0$, is the shortest vector in the lattice $L_C(\alpha)$, then u is the denominator of a convergent of the continued fraction of α and $u \leq \frac{2}{\sqrt{3}}\sqrt{C}$.*

Proof: Assume the opposite and let n be the largest index such that $q_n < u$. Then from properties of the best approximations (Theorem 8.29), it follows that $|p_n - \alpha q_n| < |w - \alpha u|$. Hence,

$$q_n^2 + C^2(p_n - \alpha q_n)^2 < u^2 + C^2(w - \alpha u)^2,$$

which is a contradiction to the minimality of $(u, C(w - \alpha u))^\tau$.

Since $\Delta(L_C(\alpha)) = C$, from Proposition 8.56 and $\gamma_2 = \frac{4}{3}$, it follows that $u \leq \frac{2}{\sqrt{3}}\sqrt{C}$. □

We conclude that the LLL algorithm can be used for obtaining good rational approximations. Unlike the continued fraction algorithm, which provides a sequence of good approximations, here for fixed C we obtain one good approximation, so for obtaining more good approximations, it is necessary to vary C.

We will now apply this idea to the problem of simultaneous Diophantine approximations.

Theorem 8.59. *Let $\alpha_1, \ldots, \alpha_n$ be real numbers and $Q > 1$ an integer. There is a polynomial time algorithm which finds integers $q, p_1, \ldots, p_n$ such that*

$$1 \leq q \leq 2^{n/4}Q^n \quad \textit{and} \quad |\alpha_i q - p_i| \leq \frac{\sqrt{5} \cdot 2^{(n-4)/4}}{Q}, \ i = 1, \ldots, n. \tag{8.46}$$

Proof: For a real number x, by $\lfloor x \rceil$ we denote the nearest integer to x. Let $C = Q^{n+1}$. Consider the matrix

$$B = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ -\lfloor C\alpha_1 \rceil & C & 0 & \cdots & 0 \\ -\lfloor C\alpha_2 \rceil & 0 & C & \cdots & 0 \\ \vdots & 0 & 0 & \ddots & \vdots \\ -\lfloor C\alpha_n \rceil & 0 & 0 & \cdots & C \end{pmatrix}$$

and the lattice L generated by its columns. The dimension of the lattice is $n + 1$, and its volume is C^n. By Lemma 8.57.3), the LLL algorithm finds a vector

$$r = (q, p_1, \ldots, p_n)^\tau$$

from $\mathbb{Z}^{n+1}$, such that $v = Br$ is a vector from L whose norm is

$$\leq \Lambda := 2^{n/4} \cdot C^{n/(n+1)} = 2^{n/4} Q^n.$$

The components of the vector v are $v_1 = q$, $v_{i+1} = Cp_i - q\lfloor C\alpha_i \rceil$, $i = 1, \ldots, n$. We have

$$q^2 + \sum_{i=1}^{n} (Cp_i - q\lfloor C\alpha_i \rceil)^2 \leq \Lambda^2.$$

Thus, $q \leq \Lambda$ and $\max_{1 \leq i \leq n} |Cp_i - q\lfloor C\alpha_i \rceil| \leq \Lambda$. Since $|Cp_i - q\lfloor C\alpha_i \rceil| \geq C|p_i - q\alpha_i| - q/2$, we obtain

$$|p_i - q\alpha_i| \leq C^{-1}(|Cp_i - q\lfloor C\alpha_i \rceil| + q/2).$$

Finally, let us use this simple fact: for real numbers x, y,

$$2x + y \leq \sqrt{5(x^2 + y^2)}. \tag{8.47}$$

Indeed, by squaring, we obtain $(x - 2y)^2 \geq 0$, which is obviously true. If we apply (8.47) to $x = |Cp_i - q\lfloor C\alpha_i \rceil|$, $y = q$, we obtain

$$|p_i - q\alpha_i| \leq C^{-1} \cdot \sqrt{5(|Cp_i - q\lfloor C\alpha_i \rceil|^2 + q^2)}/2 \leq \frac{\sqrt{5}}{2C}\Lambda = \frac{\sqrt{5}}{2} 2^{n/4} Q^{-1}. \qquad \square$$

Example 8.9. Let $\alpha_1 = \sqrt{2}$, $\alpha_2 = \sqrt{3}$, $\alpha_3 = \sqrt{5}$. Let us take $Q = 1000$ and apply the algorithm from Theorem 8.59. We form a matrix B as in the proof of the theorem and by using PARI, we calculate $\texttt{qflll}(B)$. From the first column of the obtained matrix, we read numbers $q = 118452669$, $p_1 = 167517371$, $p_2 = 205166041$, $p_3 = 264868220$. We get $q < 1.2 \cdot Q^3$, $|q\sqrt{2} - p_1| < 0.91 \cdot Q^{-1}$, $|q\sqrt{3} - p_2| < 0.14 \cdot Q^{-1}$, $|q\sqrt{5} - p_3| < 0.29 \cdot Q^{-1}$,

$$\max\left(\left|\sqrt{2} - \frac{p_1}{q}\right|, \left|\sqrt{3} - \frac{p_2}{q}\right|, \left|\sqrt{5} - \frac{p_3}{q}\right|\right) < q^{-4/3}.$$

Hence, the obtained simultaneous rational approximations are even better than Theorem 8.59 guarantees, and they satisfy inequality (8.37) from the corollary of Dirichlet's theorem on simultaneous approximations without any additional factors.

8.10 Exercises

1. Find all solutions of the inequality

$$\left|\sqrt{3} - \frac{p}{q}\right| < \frac{1}{q^2},$$

where p and q are relatively prime positive integers and $q < 100$.