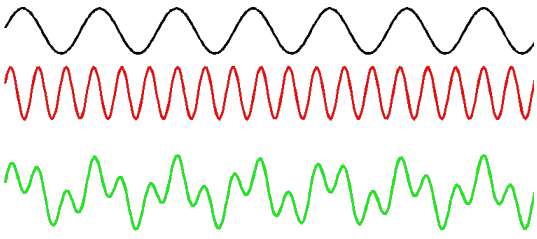
THE ABEL PRIZE 2021

Lovász-Lenstra-Lenstra lattice reduction algorithm

Almost 200 years ago, the French mathematician Jean Baptiste Joseph Fourier proved that any continuous function could be written as an infinite sum of sine and cosine waves. His result has far-reaching implications e.g. for recording and reproducing sound. A pure sine wave can be converted into sound by a loudspeaker and the sounds from orchestra instruments can be considered to be a superposition of sine waves of a fundamental frequency and integer multiples of that frequency. Each frequency contributes to the sound with an individual amplitude. The challenge when recording the sound is to separate the spectrum of frequencies.



In a mathematical framework we can consider each frequency as a separate coordinate in a huge vector space. Introducing an appropriate inner product the frequency coordinate system constitutes an orthogonal basis for the space of sine waves of different frequencies. The decomposition procedure is much less time consuming in an orthogonal basis compared to an arbitrary basis. This is a prominent example of why it is relevant to find orthogonal bases.

An efficient algorithm for producing orthogonal bases is the Gram-Schmidt process.

Lemma. Let $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\} \subset \mathbb{R}^n$ be a basis for a vector space V, and let

$$\mu_{ij} = \frac{\langle \mathbf{b}_i, \mathbf{b}_j^* \rangle}{\|\mathbf{b}_j^*\|^2}, \quad 1 \leq j < i \leq n$$

These numbers are called the Gram-Schmidt coefficients of the basis $\mathcal{B}$. They carry information about the orthogonal projection of one basis vector along another. Let

$$\mathbf{b}_k^* = \mathbf{b}_k - \sum_{i=1}^{k-1} \mu_{ki} \mathbf{b}_i^*$$

Then $\mathcal{B}^* = \{\mathbf{b}_1^*, \mathbf{b}_2^*, \ldots, \mathbf{b}_n^*\}$ is an orthogonal basis for $\mathbb{R}^n$.

Given an orthogonal basis of a vector space, it is easily transformed into an orthonormal basis, i.e. a basis that is orthogonal and such that each basis vector has length 1.

Computer science is by nature discrete and not continous. We are therefor forced to to work over a lattice $\mathbb{Z}^n$, rather than the vector space $\mathbb{R}^n$. It is still very useful to have an orthogonal basis at hand, and also ”short” basis vectors. The definition of a ”short” vector is rather vague, it refers to the ordinary length of the vector, and a vector of a lattice is considered to be short if its length is close to the minimal length of any vector in the lattice. As an example, consider the lattice generated by the two ”long” vectors $(6386, 51)$

and $(71999, 575)$. Another basis for the same lattice consists of the "short" vectors $(1, 0)$ and $(0, 1)$.

The problem for lattices versus vector spaces is the following obstructions:

i) It is not obvious that there exists an orthogonal basis for the lattice,

Ii) It can be rather difficult to find "short" vectors in the lattice.

A solution to this problemt is to produce a Lovász-Lenstra-Lenstra-reduced basis (or a LLL-reduced basis for short):

Definition. A basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ for the lattice $\mathcal{L}$ is LLL-reduced if there exists a parameter $\delta \in (0.25, 1]$ such that the following holds:

(i) For $1 \leq j < i \leq n$: $|\mu_{ij}| \leq 0.5$ (size-reduction)

(ii) For $k = 2, 3, \dots, n$: $\delta \|\mathbf{b}_{k-1}^*\|^2 \leq \|\mathbf{b}_k^* + \mu_{k,k-1}\mathbf{b}_{k-1}^*\|^2$ (Lovász condition)

where $\mathcal{B}^*$ denotes the Gram-Schmidt basis.

Using the ordinary euclidean inner product we can write

$$\langle \mathbf{b}_i, \mathbf{b}_j^* \rangle = \|\mathbf{b}_i\| \, \|\mathbf{b}_j^*\| \, \cos(\theta)$$

where θ is the angle between $\mathbf{b}_i$ and $\mathbf{b}_j^*$. The first condition then looks like

$$\|\mathbf{b}_j^*\| \, |\cos(\theta)| \leq \frac{\|\mathbf{b}_i\|}{2}$$

The inequality is satisfied if $\mathbf{b}_i$ and $\mathbf{b}_j$ are close to being orthogonal, i.e. $\cos(\theta)$ is close to 0, or if the basis vectors are ordered by length. For an orthogonal basis the Lovász condition says that

$$\delta^{n-1}\|\mathbf{b}_1^*\|^2 \leq \delta^{n-2}\|\mathbf{b}_2^*\|^2 \leq \cdots \leq \delta\|\mathbf{b}_{n-1}^*\|^2 \leq \|\mathbf{b}_n^*\|^2$$

thus in any case we get the length ordering of basis vectors.

In a paper from 1972 the Abel Prize Laureate László Lovász together with the Lenstra brothers, Arjen and Hendrik, invented the so-called LLL algorithm. The LLL algorithm is designed to produce an LLL reduced basis with an arbitrary basis for the lattice as input, and it has two main components; length reduction and basis vector swapping. Length reduction is performed by a Gram-Schmidt-type process, and swapping is needed to keep the rough length ordering of the basis vectors intact.

As an illustration of the algorithm, consider the lattice $\mathcal{L} \simeq \mathbb{Z}^3$ spanned by the three vectors

$$\mathbf{b}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \quad \mathbf{b}_2 = \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix} \quad \mathbf{b}_3 = \begin{pmatrix} 3 \\ 5 \\ 6 \end{pmatrix}$$

The Euclidean length of the three vectors is $\sqrt{3}$, $\sqrt{5}$ and $\sqrt{70}$. After performing the LLL algorithm we end up with the basis

$$\mathbf{v}_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad \mathbf{v}_3 = \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix}$$

where $\|\mathbf{v}_1\| = 1$, $\|\mathbf{v}_2\| = \sqrt{2}$ and $\|\mathbf{v}_3\| = \sqrt{5}$. The Gram-Schmidt coefficients of the resulting basis are $\mu_{21} = \mu_{31} = 0$, $\mu_{32} = \frac{1}{2}$.

The LLL algorithm has several applications, in particular within cryptography. But another interesting application is the Mertens conjecture. The Mertens conjecture was stated towards the end of the nineteenth century by the Polish mathematician Franz Mertens. The conjecture basically concerns the distribution of primes, and the most prominent consequence of the conjecture is the Riemann hypothesis. Unfortunately, the conjecture fails to be true, and attack on the Riemann hypothesis along the Mertens pathway did therefore not succeed.

The essential ingredient of the Mertens Conjecture is the Möbius function. The Möbius function μ was introduced by the German mathematician August Ferdinand Möbius in 1832. It is defined as follows:

- If n contains a square, then $\mu(n) = 0$.
- If n is a square-free number with r prime factors, then $\mu(n) = (-1)^r$.

Thus for some small numbers we have $\mu(1) = 1$, $\mu(2) = \mu(3) = \mu(5) = -1$, $\mu(4) = 0$ and $\mu(6) = 1$.

Whenever given a sequence of numbers we can collect them in a so-called Dirichelet series;

$$\begin{aligned} \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s} &= 1 - \frac{1}{2^s} - \frac{1}{3^s} - \frac{1}{5^s} + \frac{1}{6^s} - \frac{1}{7^s} \cdots \\ &= (1 - \frac{1}{2^s})(1 - \frac{1}{3^s})(1 - \frac{1}{5^s}) \cdots \end{aligned}$$

where s is a complex variable. There is a close relation between the Dirichelet series of the Möbius function and the Riemann zeta function

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = 1 + \frac{1}{2^s} + \frac{1}{3^s} + \frac{1}{4^s} + \dots$$

Multiplying the Riemann zeta function by $\frac{1}{2^s}$ and subtracting, we get

$$\begin{aligned} (1 - \frac{1}{2^s})\zeta(s) &= (1 + \frac{1}{2^s} + \frac{1}{3^s} + \frac{1}{4^s} + \dots) \\ &\quad - (\frac{1}{2^s} + \frac{1}{4^s} + \frac{1}{6^s} + \dots) \\ &= 1 + \frac{1}{3^s} + \frac{1}{5^s} + \frac{1}{7^s} \cdots \end{aligned}$$

We do the same for all primes and get

$$(1-\frac{1}{2^s})(1-\frac{1}{3^s})(1-\frac{1}{5^s})\cdots = \frac{1}{\zeta(s)}$$

Thus the Dirichelet series of the Möbius function coincides with the reciprocal of the Riemann zeta function.

$$\sum_{n=1}^{\infty}\frac{\mu(n)}{n^s} = \frac{1}{\zeta(s)}$$

The Mertens Conjecture was stated by Thomas Joannes Stieltjes, in an 1885 letter to Charles Hermite and again in print by Franz Mertens in 1897. The Mertens Conjecture gives an upper bound when adding up the Möbius function;

Conjecture. We have

$$M(x) = \sum_{n<x}\mu(n) \leq \sqrt{x}$$

for all $x \geq 1$.

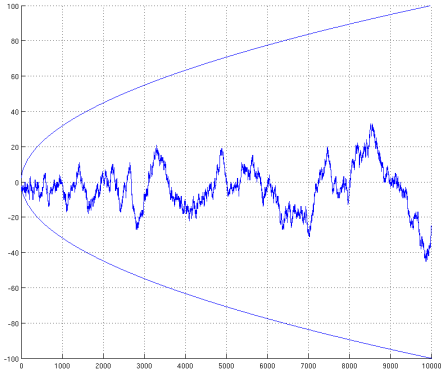
Notice that since $M(x)$ is constant on each interval $[n, n+1)$ we have

$$\begin{aligned}\frac{1}{\zeta(s)} = \sum_{n=1}^{\infty}\frac{\mu(n)}{n^s} &= \sum_{n=1}^{\infty}\frac{M(n)-M(n-1)}{n^s}\\ &= \sum_{n=1}^{\infty}M(n)\frac{1}{n^s} - \sum_{n=0}^{\infty}M(n)\frac{1}{(n+1)^s}\\ &= \sum_{n=1}^{\infty}M(n)\left(\frac{1}{n^s}-\frac{1}{(n+1)^s}\right)\\ &= \sum_{n=1}^{\infty}M(n)\cdot\int_n^{n+1}\frac{s\,dx}{x^{s+1}}\\ &= s\int_0^{\infty}\frac{M(x)\,dx}{x^{s+1}}\end{aligned}$$

Now suppose that the Mertens conjecture was true, i.e. $M(x) \leq \sqrt{x}$. Then we have

$$\frac{1}{\zeta(s)} \leq s\int_0^{\infty}\frac{\sqrt{x}\,dx}{x^{s+1}} = s\int_0^{\infty}\frac{dx}{x^{s+\frac{1}{2}}}$$

The last integral defines an analytic function for $\Re(s) > \frac{1}{2}$, and this would give an analytic continuation of $\frac{1}{\zeta(s)}$ to $\Re(s) > \frac{1}{2}$. In particular, this would imply that $\zeta(s)$ has no zeros in $\Re(s) > \frac{1}{2}$, which is exactly the statement of the Riemann hypothesis. Thus, if the Mertens Conjecture was proven to be true, the same would be the case for the Riemann Hypothesis. Unfortunately, the Mertens Conjecture is not true. It was disproved by Andrew Odlyzko and Herman te Riele in 1985. Nevertheless, the Mertens Conjecture is a striking example of a mathematical conjecture proven false despite a large amount of computational evidence in its favour, as illustrated in the figure below. The outer curve is the function $f(x) = \pm\sqrt{x}$ and the inner zig-zag curve is the Mertens function $M(x)$. The conjecture states that the zig-zag curve stays within the outer curve for all $x \in \mathbb{R}_+$.



A crucial point in the proof of Odlyzko and te Riele was an application of the lattice basis reduction algorithm of Lenstra, Lenstra and Lovász. Using the LLL algorithm they were able to show that

$$\limsup_{x\to\infty}\frac{M(x)}{\sqrt{x}} > 1.06$$

and

$$\liminf_{x\to\infty}\frac{M(x)}{\sqrt{x}} < -1.009$$

which contradicts the inequality of the conjecture. In their argument they do not give an exact value for which the conjecture fails, but it was later shown that the first counterexample appears in the gap between 10^{16} and 10^{40}.