Approximating irrational numbers by rational ones

Bill Casselman
University of British Columbia
cass@math.ubc.ca

In this note I’ll explain the geometry of approximation by continued fractions.
Suppose $\lambda$ to be a positive, irrational, real number. Given a positive integer $q$, the number $\lambda$ lies in one of the half-open intervals $[p/q, (p + 1)/q)$, in which case

$$\left| \frac{\lambda - p}{q} \right| < \frac{1}{q}.$$  

The inequality above is equivalent to $|p - q\lambda| < 1$.

Given $q$, finding $p$ is simple since

$$|q\lambda| \leq q\lambda < |q\lambda| + 1.$$  

This can be pictured—the claim is just that the line $y = \lambda x$ meets the line $x = q$ at a distance less than 1 from the point $(q, p)$. The figure below illustrates how things look when $\lambda$ is the golden ratio $\tau = (1 + \sqrt{5})/2 \sim 1.618$.

Of course we can do slightly better without extra work, getting a gap of $1/2q$ rather than $1/q$. This is insignificant. But the figure also shows that the way in which lattice points approach the line $y = \tau x$ is rather erratic. There are evidently some which are very close to the line, meaning that some fractions approximate $\tau$ very closely. The following well known result tells us that we can do arbitrarily well:

1. Proposition. There exist points of the lattice arbitrarily close to the line $y = \lambda x$.

Proof. It must be shown that for any $N > 0$ there exists $m, n$ such that $|n\lambda - m| < 1/N$.  

How lattice points approximate the line $y = \tau x$
Plot the $N + 1$ numbers $n \lambda - \lfloor n \lambda \rfloor$ for $1 \leq n \leq N + 1$. Because there are $N$ intervals $[i/N, (i + 1)/N]$ in $[0, 1]$ and $N + 1$ of these numbers, at least two of the numbers must lie in the same interval, say for $n_1 \lambda$ and $n_2 \lambda$ with $n_2 > n_1$. But then there exists some integer $m$ such that

$$|(n_2 \lambda - n_1 \lambda) - m| = |(n_2 - n_1)\lambda - m| < 1/N.$$  

There is a very simple if inefficient procedure for finding good approximations, just scanning horizontally from one approximation until you find a better one. However, we can do much better by using continued fractions. I’ll present here an elegant graphical version of this method due implicitly to the nineteenth century English mathematician Henry J. Smith (in a casual observation at the end of [Smith:1876]) and explicitly to Felix Klein.

Let $\lambda^+$ be the set of all lattice points in the positive quadrant above the line $y = \lambda x$, and $\lambda^-$ be that of all lattice points in the same quadrant below it. Let $C^\pm$ be the corresponding convex hulls.

The vertices of the sets $C^\pm$ will certainly give good approximations to $\lambda$, and it turns out that there is a very efficient way to find them. The starting point is the very elementary observation:

- $(0, 1)$ is a vertex of $C^+$;
- $(1, 0)$ and $(1, \lfloor \lambda \rfloor)$ are vertices of $C^-$.

The coordinates here are, of course, in the standard rectangular system.
But now I change the basis of the lattice $\mathbb{Z}^2$. To make notation simpler, set $\lambda_0 = \lambda$ and $\ell_0 = [\lambda]$. The old basis is $u_0 = (1, 0)$ and $v_0 = (0, 1)$, while the new one is

\[
\begin{align*}
u_1 &= v_0 = (0, 1) \\
v_1 &= u_0 + \ell_0 v_0 = (1, \ell_0).
\end{align*}
\]

In other words, I shear the original basis and reverse orientation as well. With this new basis, we have a new coordinate system $(x_1, y_1)$. The change of coordinates is

\[
\begin{align*}
x_0 &= y_1 \\
y_0 &= x_1 + \ell_0 y_1.
\end{align*}
\]

Hence:

- The line $y_0 = \lambda_0 x_0$ is now the line

\[
y_1 = \frac{x_1}{\lambda_0 - \ell_0} = \lambda_1 x_1.
\]

if $\lambda_1 = 1/(\lambda_0 - \ell_0)$.

The next basic observation is:

*The extremal points of each of the regions $\lambda^\pm$ in the new coordinate system lying along the line are the same as those of the old, except that we have lost the point that was $(1, 0)$ in the old system.*
In other words, we are in essentially the same situation as when we started out.

This leads to an infinite inductive process. We may continue on forever to find new extremal points, one at a time, as well as edges connecting them to one of the extremal points already found. Expressing the points we find in the original coordinate system, we find the approximations we are looking for, as in the classical form of continued fractions.

We get in this way a sequence of bases \((u_n, v_n)\):

\[
\begin{align*}
  u_0 &= (1, 0) \\
  v_0 &= (0, 1) \\
  \lambda_0 &= \lambda \\
  \ell_0 &= \lfloor \lambda_0 \rfloor \\
  u_1 &= (0, 1) \\
  v_1 &= (1, \ell_0) \\
  \lambda_1 &= 1/(\lambda_0 - \ell_0) \\
  \ell_1 &= \lfloor \lambda_1 \rfloor \\
  \vdots \\
  u_n &= v_{n-1} \\
  v_n &= u_{n-1} + \ell_{n-1} v_{n-1} \\
  \lambda_n &= 1/(\lambda_{n-1} - \ell_{n-1}) \\
  \ell_n &= \lfloor \lambda_n \rfloor.
\end{align*}
\]

The equation relating \(\lambda_n\) and \(\lambda_{n+1}\) can be reformulated as

\[
\lambda_n = \ell_n + \frac{1}{\lambda_{n+1}}. \tag{2}
\]

We then deduce by induction a sequence of equations:

\[
\begin{align*}
  \lambda &= \lambda_0 \\
  &= \ell_0 + \frac{1}{\ell_1 + \lambda_1} \\
  &= \ell_0 + \frac{1}{\ell_1 + \frac{1}{\ell_2 + \lambda_2}} \\
  &= \ell_0 + \frac{1}{\ell_1 + \frac{1}{\ell_2 + \frac{1}{\ell_3 + \lambda_3}}} \\
  \vdots
\end{align*}
\]

\[
\begin{align*}
  \lambda &= \ell_0 + \frac{1}{\ell_1 + \frac{1}{\ell_2 + \frac{1}{\ell_3 + \lambda_3}}} \\
  &\vdots
\end{align*}
\]
leading to the approximations

\[ \lambda \sim \ell_0 \]
\[ \sim \ell_0 + \frac{1}{\ell_1} \]
\[ \sim \ell_0 + \frac{1}{\ell_1 + \frac{1}{\ell_2}} \]
\[ \sim \ell_0 + \frac{1}{\ell_1 + \frac{1}{\ell_2 + \frac{1}{\ell_3}}} \]
\[ \ldots \]

This explains the term ‘continued fractions’. To avoid typesetting problems, this is often expressed as

\[ \ell_0 + \frac{1}{\ell_1 + \frac{1}{\ell_2 + \frac{1}{\ell_3 + \ldots}}} . \]

The recursion formulas can be encapsulated in terms of matrices

\[ \Gamma_n = \begin{bmatrix} u_n & v_n \end{bmatrix} = \text{(say)} \begin{bmatrix} q_{n-1} & q_n \\ p_{n-1} & p_n \end{bmatrix} \]

whose columns are the vectors in the \( n \)-th basis:

\[ \Gamma_0 = \begin{bmatrix} u_0 & v_0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \]
\[ \Gamma_1 = \begin{bmatrix} u_1 & v_1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & \ell_0 \end{bmatrix} \]
\[ \ldots \]
\[ \Gamma_n = \begin{bmatrix} u_n & v_n \end{bmatrix} = \Gamma_{n-1} \begin{bmatrix} 0 & 1 \\ 1 & \ell_{n-1} \end{bmatrix} . \]

Induction shows that \( \det(\Gamma_n) = (-1)^n \), which implies an agreeable fact:

**4. Proposition.** *In either column of \( \Gamma_n \), the coordinates are relatively prime.*

The formulas (3) imply that

\[ \lambda = \frac{a_n \lambda_n + b_n}{c_n \lambda_n + d_n} \]

for suitable coefficients. What are they? To start with:

\[ \lambda = \frac{1 \cdot \lambda_0 + 0}{0 \cdot \lambda_0 + 1} . \]
We can proceed by induction. If we are given (5), we can apply (2) with \( n - 1 \) instead of \( n \) to get

\[
\lambda = \frac{a_{n-1} \left( \ell_{n-1} + \frac{1}{\lambda_{n-1}} \right) + b_{n-1}}{c_{n-1} \left( \ell_{n-1} + \frac{1}{\lambda_{n-1}} \right) + d_{n-1}} = \frac{(a_{n-1} \ell_{n-1} + b_{n-1}) \lambda_n + a_{n-1}}{(c_{n-1} \ell_{n-1} + d_{n-1}) \lambda_n + c_{n-1}}
\]

so

\[
\begin{align*}
a_n &= a_{n-1} \ell_{n-1} + b_{n-1} \\
b_n &= a_{n-1} \\
c_n &= c_{n-1} \ell_{n-1} + d_{n-1} \\
d_n &= c_{n-1}
\end{align*}
\]

\[
\begin{bmatrix} a_n & b_n \\ c_n & d_n \end{bmatrix} = \begin{bmatrix} a_{n-1} & b_{n-1} \\ c_{n-1} & d_{n-1} \end{bmatrix} \begin{bmatrix} \ell_{n-1} & 1 \\ 1 & 0 \end{bmatrix}.
\]

In other words

\[
\begin{bmatrix} a_n & b_n \\ c_n & d_n \end{bmatrix} = w \cdot \Gamma_n \cdot w^{-1} \quad \left( w = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \right).
\]

Hence:

6. Proposition. For every \( n \geq 0 \)

\[
\lambda = \frac{p_n \lambda_n + p_{n-1}}{q_n \lambda_n + q_{n-1}}
\]

The occurrence of the matrix \( w \) is natural, since fractions are written as \( p/q \), while coordinates as \( (q, p) \).

The corresponding \( n \)-th approximation to \( \lambda \) is

\[
\tilde{\lambda}_n = \frac{p_n \ell_n + p_{n-1}}{q_n \ell_n + q_{n-1}} = \frac{p_{n+1}}{q_{n+1}}
\]

and then

\[
\lambda - \tilde{\lambda}_n = \frac{p_n \lambda_n + p_{n-1}}{q_n \lambda_n + q_{n-1}} - \frac{p_n \ell_n + p_{n-1}}{q_n \ell_n + q_{n-1}} = \frac{(p_n q_n - p_{n-1} q_n) (\lambda_n - \ell_n)}{(q_n \lambda_n + q_{n-1})(q_n \ell_n + q_{n-1})}.
\]

If \( v_n = (q_n, p_n) \), the \( q_n \) make up a strictly increasing sequence; the determinant of \( \Gamma_n \) is \( \pm 1 \); and \( |\lambda_n - \ell_n| < 1 \).

Therefore:

8. Proposition. For every \( n \geq 0 \)

\[
| \lambda - \frac{p_n}{q_n} | < \frac{1}{q_n^2}.
\]

Examples. (1) If \( \lambda \) is rational, finding the continued fraction of \( \lambda \) amounts to carrying out Euclid’s algorithm for greatest common denominator.
If $\lambda$ is the root of a quadratic polynomial with rational coefficients, its continued fraction is eventually repetitive. For example, say $\lambda = \lambda_0 = \tau$. Then

$$\lambda_0 = \frac{1 + \sqrt{5}}{2}$$

$$\lambda_1 = \frac{1}{\lambda_0 - 1} = \frac{1}{(-1 + \sqrt{5}/2)} = \frac{2(-1 - \sqrt{5})}{1 - 5}$$

so that $\ell_n = 1$ for every $n$. We get the sequence of matrices $\Gamma_{2n}$

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}^2, \quad \begin{bmatrix} 2 & 3 \\ 3 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}, \quad \begin{bmatrix} 5 & 8 \\ 8 & 13 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}, \quad \cdots$$

The earliest systematic reference to this phenomenon that I know of is [Euler:1744].

Of course the sequence for a rational number is also repetitive in some sense. One way to interpret this is that if $\lambda$ is rational, we eventually pass off to $1/0 = \infty$, which is a fixed point of the process.

(3) Other examples for which the behaviour of the sequence $(\ell_n)$ is regular or even predictable are rare. [Euler:1744] finds the sequence of the $\ell_n$ for several numbers related to $e$. For $e$ itself it is

$$2, 1, 2, 1, 1, 4, 1, 1, 6, 1, 1, 8, \ldots$$

He finds this by an analysis (in §§28 ff.) of solutions of Riccati’s differential equation.

References


There are many places to find out anything you want to know about continued fractions. This is more imaginative than most.