Lecture 2: Continued fractions, rational approximations

Algorithmic Number Theory (Fall 2014)
Rutgers University
Swastik Kopparty
Scribe: Cole Franks

1 Continued Fractions

We begin by calculating the continued fraction representation of a real number. Let α be a real number. Then

$$\begin{aligned} \alpha &= \lfloor \alpha \rfloor + \frac{1}{\alpha_1}, \\ \alpha_1 &= \lfloor \alpha_1 \rfloor + \frac{1}{\alpha_2}, \\ &\vdots \\ \alpha_{n-1} &= \lfloor \alpha_{n-1} \rfloor + \frac{1}{\alpha_n} \end{aligned}$$

where α_i is greater than one for $i \in \{1, \ldots, n\}$. Let $a_0 := \lfloor \alpha \rfloor$, and $a_i := \lfloor \alpha_i \rfloor$. At some point α_N may be an integer, in which case this process terminates and α is rational. In that case, we simply define $a_n = 0$ for all $n > N$, and α_n is undefined for $n > N$. Note that $a_n \geq 1$ for all $1 \leq n \leq N$. If α is irrational, the process will never terminate, so we don't have to worry about the two previous lines. Note that $a_i, \alpha_i \geq 0$ for all $i \geq 1$. By plugging each equation into the one above, we see that

$$\alpha = a_0 + \cfrac{1}{a_1 + \cfrac{1}{\ddots \cfrac{\vdots}{a_{n-1} + \cfrac{1}{\alpha_n}}}}.$$

The n^{th} convergent of α, provided $n \leq N$ (if N exists), is given by

$$c_n = a_0 + \cfrac{1}{a_1 + \cfrac{1}{\ddots \cfrac{\vdots}{a_{n-1} + \cfrac{1}{a_n}.}}}$$

If $n \geq N$, then the n^{th} convergent of α is simply

$$c_n = a_0 + \cfrac{1}{a_1 + \cfrac{1}{\ddots \cfrac{\vdots}{a_{N-1} + \cfrac{1}{a_N}.}}} = c_N$$

The list $[a_0, \dots]$ is the *continued fraction representation* of α.

Lemma 1. *The convergents are given by* $c_n = \frac{p_n}{q_n}$ *where*

$$\begin{bmatrix} p_{n+1} & p_n \\ q_{n+1} & q_n \end{bmatrix} = \begin{bmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{bmatrix} \begin{bmatrix} a_{n+1} & 1 \\ 1 & 0 \end{bmatrix}$$

or

$$\begin{aligned} p_{n+1} &= p_n a_{n+1} + p_{n-1} \\ q_{n+1} &= q_n a_{n+1} + q_{n-1} \end{aligned}$$

with $p_0 = a_0$, $q_0 = 1$, $p_1 = a_0 a_1 + 1$, $q_1 = a_1$.

Note that for all $i \geq 0$, $q_i > 0$. In particular, q_i is strictly monotonically increasing until $a_{i+1} = 0$, or when $i = N$.

Proof. First we claim that for any $n \geq 1$ and a formal variable Z we have

$$\frac{p_n Z + p_{n-1}}{q_n Z + q_{n-1}} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{\ddots \cfrac{\vdots}{a_n + \cfrac{1}{Z},}}} \tag{1}$$

which we prove by induction. The base case $n = 1$ is an exercise. Let $n \geq 1$. From the inductive hypothesis with $Z \leftarrow a_{n+1} + \frac{1}{Z}$,

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{\ddots \cfrac{\vdots}{a_{n+1} + \cfrac{1}{Z},}}} = \frac{p_n \left(a_{n+1} + \frac{1}{Z}\right) + p_{n-1}}{q_n \left(a_{n+1} + \frac{1}{Z}\right) + q_{n-1}} = \frac{p_n \left(a_{n+1} Z + 1\right) + p_{n-1} Z}{q_n \left(a_{n+1} Z + 1\right) + q_{n-1} Z}$$

$$= \frac{p_n a_{n+1} Z + p_n + p_{n-1} Z}{q_n a_{n+1} Z + q_n + q_{n-1} Z} = \frac{p_{n+1} Z + p_n}{q_{n+1} Z + q_n}.$$

If we substitute $Z = a_{n+1}$ in (**??**), we get that $c_{n+1} = p_{n+1}/q_{n+1}$ as desired. □

Observation 2. *Equation (**??**) is valid even if $n > N$, for the proof did not use $n \leq N$.*

Note that

$$\det\left(\begin{bmatrix} p_1 & p_0 \\ q_1 & q_0 \end{bmatrix}\right) = -\det\left(\begin{bmatrix} a_{n+1} & 1 \\ 1 & 0 \end{bmatrix}\right) = 1,$$

so

$$\det\left(\begin{bmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{bmatrix}\right) = (-1)^{n-1}.$$

This implies that p_n, q_n are relatively prime (because the ideal they generate in $\mathbb{Z}$ contains 1), and so the convergent $\frac{p_n}{q_n}$ is already in lowest terms.

2 Approximation

Now we move to the approximation of real numbers by rational numbers. Our aim is to use the lowest denominator rational number possible and still get a nice approximation.

How good a rational approximation can one get to a given real number α? One trivial rational approximation to α is a number $\frac{a}{q}$ with

$$\left|\alpha - \frac{a}{q}\right| \leq \frac{1}{2q},$$

(for any q we can simply choose $a \in \mathbb{Z}$ at a distance at most $1/2$ from $q\alpha$). The following theorem shows that every α has a pretty good rational approximation:

Theorem 3 (Dirichlet's Theorem)**.** *For every $\alpha \in \mathbb{R}$, either α is rational, or else there are infinitely many distinct rational numbers $\frac{a}{q}$ such that*

$$\left|\alpha - \frac{a}{q}\right| < \frac{1}{q^2}.$$

Observation 4.

We prove Theorem **??** by proving the following stronger theorem:

Theorem 5. *For all $Q \in \mathbb{N}$, there is a rational number $\frac{a}{q}$ such that $q \leq Q$ and*

$$\left|\alpha - \frac{a}{q}\right| < \frac{1}{qQ}.$$

Lemma 6. *Theorem **??** implies Theorem **??**.*

*Proof of Lemma **??**.* Take some $Q_1 \in \mathbb{N}$; by Theorem **??** we can find $\frac{a_1}{q_1}$ with

$$\left|\alpha - \frac{a_1}{q_1}\right| < \frac{1}{q_1 Q_1} \leq \frac{1}{q_1^2}.$$

If $\frac{a_1}{q_1} = \alpha$, then α is rational, and so we are done with Theorem **??** in this case. If not, i.e. $\left|\alpha - \frac{a_1}{q_1}\right| > 0$, then we can take a new Q_2 so large that $\left|\alpha - \frac{a_1}{q_1}\right| > \frac{1}{Q_2}$. Let $\frac{a_2}{q_2}$ be the pair guaranteed by Theorem **??** for Q_2. Then

$$\left|\alpha - \frac{a_2}{q_2}\right| < \frac{1}{q_2 Q_2} < \frac{\left|\alpha - \frac{a_1}{q_1}\right|}{q_2} \leq \left|\alpha - \frac{a_1}{q_1}\right|,$$

so $\frac{a_2}{q_2}$ is distinct from $\frac{a_1}{q_1}$. Furthermore,

$$\left|\alpha - \frac{a_2}{q_2}\right| < \frac{1}{q_2 Q_2} \leq \frac{1}{q_2^2}.$$

Continuing this process, we either find that α is rational, or else we find infinitely many $\frac{a}{q}$ as required in Theorem **??**. □

This proof, due to Dirichlet, was possibly the first appearance of the Dirichlet Pigeon Hole Principle.

Proof of Theorem **??**. If we show that there is some q with $q\alpha$ closer than $\frac{1}{Q}$ to some integer, then we can use that integer as our a and be done. Consider $\{0 \mod 1, \alpha \mod 1, \dots Q\alpha \mod 1\}$. This is an $Q+1$ element set lying in the interval $[0,1)$. There must be two in one of the Q intervals $\left[\frac{i-1}{Q}, \frac{i}{Q}\right)$, $i \in \{1, \dots, Q\}$. Hence there exist r, s such that $0 \leq s\alpha \mod 1 - r\alpha \mod 1 = (s-r)\alpha \mod 1 < \frac{1}{Q}$, viewed as an element of $[0,1)$. Then there is an integer, namely that integer directly below $(s-r)\alpha$, such that $|(s-r)\alpha - a| < \frac{1}{Q}$. Our q is $s-r$.

□

Let us now give an algorithm to construct these good approximations. Concretely, the problem we want to solve is this:

Problem 1. *Given a rational number $\alpha = \frac{r}{s}$ and an integer Q, find $\frac{a}{q}$ with $|\alpha - \frac{a}{q}|$ as small as possible subject to $q \leq Q$.*

The input size for this problem is $O(\log r + \log s + \log Q)$ bits, and we will be interested in getting an algorithm which runs in time polynomial in this.

Our algorithm will be based on continued fractions (and we will use the notation we used in that section). It is not true that the solution to our rational approximation problem will be a convergent, but convergents will help us compute the solution easily.

We begin with a lemma on the sign of $c_n - \alpha$.

Lemma 7. *If n is even, $c_n \leq \alpha$, else $c_n \geq \alpha$.*

Proof. If $n > N$, this is trivial because $c_n = \alpha$, so we assume that $n \leq N$. We proceed by induction on n (showing the result for all α at the same time).

The lemma is clearly true for $n = 0$ (since $c_0 = a_0 = \lfloor \alpha \rfloor$). Now let $1 \le n \le N-1$, and assume the lemma holds for n for every α. Then

$$c_{n+1} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{\ddots \cfrac{\vdots}{a_n + \cfrac{1}{a_{n+1}}.}}} = a_0 + \frac{1}{c'_n},$$

where c'_n is the n^{th} convergent of α_1, the number such that $\alpha = a_0 + \frac{1}{\alpha_1}$. c'_n is either at least (if n is odd) or at most α_1 (if n is even) by our inductive assumption. By similar reasoning to that showing $\alpha_n > 0$ for $n > 0$, $c'_n > 0$. Then $a_0 + \frac{1}{c'_n}$ is either at least α (if $n+1$ is even) or at most α (if $n+1$ is odd). □

Now we characterize the best rational approximation with denominator $\le Q$.

Theorem 8. *Let n be largest number such that $q_n \le Q$, and let t be the greatest integer such that $tq_n + q_{n-1} \le Q$. Then either $\frac{p_n}{q_n}$ or $\frac{tp_n+p_{n-1}}{tq_n+q_{n-1}}$ is a solution to the rational approximation problem.*

Proof. Let $L = \frac{tp_n+p_{n-1}}{tq_n+q_{n-1}}$. We will assume n is even. The proof when n is odd will be very similar. Because n is even, $\frac{p_n}{q_n} \le \alpha$ and $\frac{p_{n+1}}{q_{n+1}} \ge \alpha$. Note that if $s = a_{n+1}$, then $\frac{sp_n+p_{n-1}}{sq_n+q_{n-1}} = \frac{p_{n+1}}{q_{n+1}}$; if $s = 0$, then $\frac{sp_n+p_{n-1}}{sq_n+q_{n-1}} = \frac{p_{n-1}}{q_{n-1}}$. As s ranges from 0 to a_{n+1}, $\frac{sp_n+p_{n-1}}{sq_n+q_{n-1}}$ ranges monotonically from $\frac{p_{n-1}}{q_{n-1}}$ to $\frac{p_{n+1}}{q_{n+1}}$. This is because

$$\frac{d}{ds}\left(\frac{sp_n + p_{n-1}}{sq_n + q_{n-1}}\right) = \frac{\det\left(\begin{bmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{bmatrix}\right)}{(sq_n + q_{n-1})^2};$$

does not change sign. Hence, L must be between $\frac{p_{n+1}}{q_{n+1}}$ and $\frac{p_{n-1}}{q_{n-1}}$, both of which are at least α. Hence L is at least α. If $\frac{r}{s}$ is strictly closer to α than both c_n and L, then $\frac{r}{s} \in I = (c_n, L)$. For contradiction, suppose $0 < s \le Q$. Then

$$\left|\frac{r}{s} - \frac{p_n}{q_n}\right| \ge \frac{1}{sq_n},$$

because $\frac{r}{s} \ne \frac{p_n}{q_n}$, and similarly

$$\left|\frac{r}{s} - L\right| \ge \frac{1}{s(tq_n + q_{n-1})}.$$

Observe that $|I| = \left|\frac{p_n}{q_n} - \frac{tp_n+p_{n-1}}{tq_n+q_{n-1}}\right| = \frac{1}{q_n(tq_n+q_{n-1})}$. Then

$$|I| = \frac{1}{q_n(tq_n + q_{n-1})} \ge \frac{1}{s(tq_n + q_{n-1})} + \frac{1}{sq_n} = \frac{1}{s}\left(\frac{1}{tq_n + q_{n-1}} + \frac{1}{q_n}\right) = \frac{1}{s}\left(\frac{tq_n + q_{n-1} + q_n}{q_n(tq_n + q_{n-1})}\right).$$

So $s \ge (t+1)q_n + q_{n-1} > Q$, a contradiction. □

Observation 9. *We saw in the proof that if n is even, $\alpha \in (c_n, L)$, and if n is odd, $\alpha \in (L, c_n)$. Further, the interval has length exactly $\frac{1}{q_n(tq_n+q_{n-1})}$. Then $|c_n - \alpha| \le \frac{1}{q_n(tq_n+q_{n-1})} \le \frac{1}{q_nq_{n-1}}$.*

As we noted before, q_n increases monotonically until $n = N$. As $\frac{a}{b}$ is the best approximation with denominator at most Q for any $Q \geq b$, we cannot have $q_n > b$ for any n. If so, then $q_{n+1} > b$ as well. By Theorem **??**, either L or $\frac{p_{n+1}}{q_{n+1}}$ would at least as good an approximation as $\frac{a}{b}$ with denominator at most q_{n+1}. For any $t > 0$, $tq_{n+1} + q_n > q_{n+1}$, so $t = 0$. Then $L = \frac{p_n}{q_n}$. This is impossible, because $\frac{p_n}{q_n}$ and $\frac{p_{n+1}}{q_{n+1}}$ are reduced and therefore neither are equal to $\frac{a}{b}$. This also applies that N exists for α if and only if α is rational., i.e. the continued fraction representation of α terminates iff α is rational.

3 Running Time

It will be of use to bound p_n and q_n. We already know that $q_n \leq b$. From Remark **??**, we know that $\left|p_n - q_n \frac{a}{b}\right| \leq \frac{1}{q_{n-1}}$. Hence $|p_n| \leq \left|\frac{q_n}{b}\right| |a| + \frac{1}{q_{n-1}} = O(a)$.

We know that the n from Theorem **??** is $O(\log(Q))$ because the q_n grow at least as fast as the n^{th} entry of the Fibonacci sequence provided $n \leq N$. At each step in the computation of the continued fraction, we compute the floor of a rational number $\alpha_i = \frac{r_i}{s_i}$, which is simply the number of times the denominator of α_i goes into the numerator. This is division with remainder, which we know to take $O(\log(r_i)\log(s_i))$ operations. We know $\alpha_i = \frac{1}{\alpha_{i-1} - \lfloor \alpha_{i-1} \rfloor}$, which has denominator the remainder of r_{i-1} under division by s_{i-1} (which is strictly smaller than s_{i-1}) and numerator s_{i-1}. As $\alpha_0 = \alpha = \frac{a}{b}$, we can conclude by induction that $r_i, s_i \leq \max(a, b)$ and so the number of operations per step is certainly at most $O((\log(a) + \log(b))^2)$. Finding t can take at most $O(\log(Q)^2)$ steps, as we find it by division with remainder of $Q - q_{n-1}$ by q_{n-1}. t is at most a_{n+1}, which is at most $q_{n+1} \leq b$. $p_n, p_{n-1} = O(a)$. Taken together, finding L must take $O(poly(\log(a), \log(b), \log(Q)))$. Finding whether L or c_n is closer to α can also certainly take at most $O(poly(\log(a), \log(b), \log(Q)))$, and finding the continued fraction is also $O(poly(\log(a), \log(b), \log(Q)))$. Hence the entire running time must be $O(poly(\log(a), \log(b), \log(Q)))$.