Scribe Notes for *Algorithmic Number Theory*
Class 9—May 29, 1998

Scribes: Cara Struble and Craig Struble

Abstract

Today we discuss Sections 4.4 and 4.5 in the text, covering continuants, continued fractions, and convergents.

1 Continuants

Throughout this discussion, we use the notation from Sections 4.2 and 4.3 for the equations given by steps of the Euclidean and extended Euclidean algorithms. Recall the matrices $M_0, M_1, \ldots, M_n$ from the extended Euclidean algorithm,

$$\begin{aligned}
M_0 &= \begin{bmatrix} a_0 & 1 \\ 1 & 0 \end{bmatrix}, \\
M_1 &= \begin{bmatrix} a_0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a_1 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} a_0a_1+1 & a_0 \\ a_1 & 1 \end{bmatrix}, \\
M_2 &= \begin{bmatrix} a_0a_1+1 & a_0 \\ a_1 & 1 \end{bmatrix} \begin{bmatrix} a_2 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} a_0a_1a_2+a_0+a_2 & a_0a_1+1 \\ a_1a_2+1 & a_1 \end{bmatrix}, \\
&\vdots
\end{aligned}$$

Now, consider the matrices of the same form with entries from $\mathbb{Z}[X_0, X_1, \ldots, X_{n-1}]$. So,

$$\begin{aligned}
&\begin{bmatrix} X_i & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} X_{i+1} & 1 \\ 1 & 0 \end{bmatrix} \cdots \begin{bmatrix} X_j & 1 \\ 1 & 0 \end{bmatrix} \\
&\quad = \begin{bmatrix} f_{11}(X_i, X_{i+1}, \ldots, X_j) & f_{12}(X_i, X_{i+1}, \ldots, X_{j-1}) \\ f_{21}(X_{i+1}, X_{i+2}, \ldots, X_j) & f_{22}(X_{i+1}, X_{i+2}, \ldots, X_{j-1}) \end{bmatrix}
\end{aligned}$$

From this we see, f_{11} is a function of $j-i+1$ variables, f_{12} is a function of $j-i$ variables, f_{21} is a function of $j-i$ variables, and f_{22} is a function in $j-i-1$ variables, where f_{11}, f_{12}, f_{21}, and $f_{22} \in \mathbb{Z}[X_0, X_1, \ldots, X_{n-1}]$. In fact, polynomials in the same number of variables have the same form.

Suppose there is exactly one polynomial for each k, denoted $Q_k(X_0, \ldots, X_{k-1})$. For some small values of k, Q_k is

$$\begin{aligned}
Q_0() &= 1 \\
Q_1(X_0) &= X_0 \\
Q_2(X_0, X_1) &= X_0X_1 + 1 \\
Q_3(X_0, X_1, X_2) &= X_0X_1X_2 + X_0 + X_2.
\end{aligned}$$

These equations work for M_0, M_1, and M_2 above with the appropriate substitutions. Q_0, Q_1, Q_2, Q_3 form the base case for an inductive definition of Q_k. Suppose that we know the continuants for Q_{k-1}, Q_{k-2}, and Q_{k-3}. We can calculate Q_k with the following equations:

$$\begin{aligned}
&\begin{bmatrix} X_0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} X_1 & 1 \\ 1 & 0 \end{bmatrix}\cdots\begin{bmatrix} X_{k-1} & 1 \\ 1 & 0 \end{bmatrix} \\
= &\begin{bmatrix} Q_k(X_0,\ldots,X_{k-1}) & Q_{k-1}(X_0,\ldots,X_{k-2}) \\ Q_{k-1}(X_1,\ldots,X_{k-1}) & Q_{k-2}(X_1,\ldots,X_{k-2}) \end{bmatrix} \\
= &\begin{bmatrix} Q_{k-1}(X_0,\ldots,X_{k-2}) & Q_{k-2}(X_0,\ldots,X_{k-3}) \\ Q_{k-2}(X_1,\ldots,X_{k-2}) & Q_{k-3}(X_1,\ldots,X_{k-3}) \end{bmatrix}\begin{bmatrix} X_{k-1} & 1 \\ 1 & 0 \end{bmatrix}
\end{aligned}$$

From the matrix multiplication, we see that

$$Q_k(X_0,\ldots,X_{k-1}) = X_{k-1}Q_{k-1}(X_0,\ldots,X_{k-2}) + Q_{k-2}(X_0,\ldots,X_{k-3}).$$

Q_k is formed by the sum $X_{k-1}Q_{k-1}$ and Q_{k-2}. Note that X_{k-1} is the last variable for Q_k. For example,

$$Q_{k-1}(X_1,\ldots,X_{k-1}) = X_{k-1}Q_{k-2}(X_1,\ldots,X_{k-2}) + Q_{k-3}(X_1,\ldots,X_{k-3}).$$

The polynomial Q_k is called the k^{th} **continuant**. Continuants are used to compute each u_i in the Euclidean algorithm.

Theorem 1.1 (Theorem 4.4.5). *Suppose the Euclidean algorithm runs on input (u, v) in n steps. Then,*

$$u_i = dQ_{n-i}(a_i,\ldots,a_{n-1})$$

where $d = \gcd(u, v)$.

This is shown for u_0, but the theorem is easily proved with appropriate substitutions.

$$\begin{aligned}
\begin{bmatrix} u_0 \\ u_1 \end{bmatrix} &= M_{n-1}\begin{bmatrix} u_n \\ 0 \end{bmatrix}, \\
u_0 &= u_nQ_n(X_0,\ldots,X_{n-1})|_{X_i=a_i} \\
&= dQ_n(X_0,\ldots,X_{n-1})|_{X_i=a_i} \\
&= dQ_n(a_0,\ldots,a_{n-1}).
\end{aligned}$$

We abbreviate continuants by

$$Q_{i+1}(X_j, X_{j+1},\ldots,X_{j+i}) = Q[j, j+i]$$

Some examples of this abbreviation are

$$\begin{aligned}
Q[0,0] &= X_0 \\
Q[1,1] &= X_1 \\
Q[0,1] &= X_0X_1 + 1 \\
Q[0,-1] &= Q[1,0] \\
&= Q[2,1] \\
&= 1.
\end{aligned}$$

Now, consider the determinant of the matrix M_n.

$$
\begin{aligned}
\det\left(\begin{bmatrix} Q[0,n] & Q[0,n-1] \\ Q[1,n] & Q[1,n-1] \end{bmatrix}\right) &= \det\left(\begin{bmatrix} X_0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} X_1 & 1 \\ 1 & 0 \end{bmatrix}\cdots\begin{bmatrix} X_n & 1 \\ 1 & 0 \end{bmatrix}\right) \\
&= (-1)^{n+1} \\
\det\left(\begin{bmatrix} Q[0,n] & Q[0,n-1] \\ Q[1,n] & Q[1,n-1] \end{bmatrix}\right) &= Q[0,n]Q[1,n-1] - Q[0,n-1]Q[1,n]
\end{aligned}
$$

The following theorem follows directly.

Theorem 1.2 (Theorem 4.4.2).

$$(-1)^{n+1} = Q[0,n]Q[1,n-1] - Q[0,n-1]Q[1,n].$$

Now suppose that the matrix M_n is split at an arbitrary index i.

$$
\begin{aligned}
M_n &= \begin{bmatrix} Q[0,n] & Q[0,n-1] \\ Q[1,n] & Q[1,n-1] \end{bmatrix} \\
&= \left(\begin{bmatrix} X_0 & 1 \\ 1 & 0 \end{bmatrix}\cdots\begin{bmatrix} X_i & 1 \\ 1 & 0 \end{bmatrix}\right)\left(\begin{bmatrix} X_{i+1} & 1 \\ 1 & 0 \end{bmatrix}\cdots\begin{bmatrix} X_n & 1 \\ 1 & 0 \end{bmatrix}\right) \\
&= \begin{bmatrix} Q[0,i] & Q[0,i-1] \\ Q[1,i] & Q[1,i-1] \end{bmatrix}\begin{bmatrix} Q[i+1,n] & Q[i+1,n-1] \\ Q[i+2,n] & Q[i+2,n-1] \end{bmatrix}
\end{aligned}
$$

This leads to the following theorem.

Theorem 1.3 (Theorem 4.4.4). *For* $0 \le i \le n-1$,

$$Q[0,n] = Q[0,i]Q[i+1,n] + Q[0,i-1]Q[i+2,n].$$

2 Continued Fractions

A rational number u/v can be approximated by r/s where $|r| < |u|$ and $|s| < |v|$. By Theorem 4.4.5, we can express u/v as

$$\frac{u_0}{u_1} = \frac{u}{v} = \frac{dQ_n(a_0, a_1, \dots, a_{n-1})}{dQ_{n-1}(a_1, a_2, \dots, a_{n-1})}.$$

This gives us the rational function

$$\frac{Q[0,n-1]}{Q[1,n-1]}.$$

We can then apply Theorem 4.4.4 to obtain

$$
\begin{aligned}
\frac{Q[0,n-1]}{Q[1,n-1]} &= \frac{Q[0,0]Q[1,n-1] + Q[0,-1]Q[2,n-1]}{Q[1,n-1]} \\
&= Q[0,0] + \frac{Q[0,-1]Q[2,n-1]}{Q[1,n-1]} \\
&= X_0 + \frac{Q[2,n-1]}{Q[1,n-1]} \\
&= X_0 + \cfrac{1}{\cfrac{Q[1,n-1]}{Q[2,n-1]}}.
\end{aligned}
$$

Doing the same manipulations iteratively, we get

$$\begin{aligned}
X_0 + \cfrac{1}{\cfrac{Q[1,n-1]}{Q[2,n-1]}} &= X_0 + \cfrac{1}{X_1 + \cfrac{Q[3,n-1]}{Q[2,n-1]}} \\
&= X_0 + \cfrac{1}{X_1 + \cfrac{1}{\cfrac{Q[2,n-1]}{Q[3,n-1]}}} \\
&= X_0 + \cfrac{1}{X_1 + \cfrac{1}{X_2 + \cdots + \cfrac{1}{X_{n-1}}}}
\end{aligned}$$

This is called a **continued fraction**. We write it as $[X_0, X_1, ..., X_{n-1}]$.

Example 2.1. For the fraction $\frac{216}{183}$, $a_0 = 1, a_1 = 5, a_2 = 1, a_3 = 1, a_4 = 5$. So

$$\begin{aligned}
\frac{216}{183} &= a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{a_4}}}} \\
&= 1 + \cfrac{1}{5 + \cfrac{1}{1 + \cfrac{1}{1 + \frac{1}{5}}}}.
\end{aligned}$$

Theorem 2.2 (Theorem 4.5.1). *A real number x has a finite representation as a continued fraction if and only if x is rational.*

We are interested in approximations that look like $Q[0,i]/Q[1,i]$.

2.1 Convergents

The i^{th} convergent of u/v is p_i/q_i where

$$\begin{aligned}
p_i &= Q_{i+1}(a_0, a_1, ..., a_i) \\
q_i &= Q_i(a_1, a_2, ..., a_i).
\end{aligned}$$

Each p_i/q_i is an approximation to u/v.

Example 2.3. Using 216/183 again, we obtain these values:

$$
\begin{array}{lll}
p_0 = Q_1(a_0) = a_0 & q_0 = Q_0() = 1 & \frac{p_0}{q_0} = \frac{1}{1} \\
p_1 = Q_2(a_0a_1) = a_0a_1 + 1 & q_1 = Q_1(a_1) = a_1 & \frac{p_1}{q_1} = \frac{6}{5} \\
p_2 = Q_3(a_0a_1a_2) = a_0a_1a_2 + a_0 + a_2 & q_2 = Q_2(a_1a_2) = a_1a_2 + 1 & \frac{p_2}{q_2} = \frac{7}{6} \\
p_3 = Q_4(a_0a_1a_2a_3) = & & \\
a_0(a_1a_2a_3 + a_1 + a_3) + a_2a_3 + 1 & q_3 = Q_3(a_1a_2a_3) = a_1a_2a_3 + a_1 + a_3 & \frac{p_3}{q_3} = \frac{13}{11} \\
p_4 = Q_5(a_0a_1a_2a_3a_4) = \frac{u_0}{d} = 72 & q_4 = Q_4(a_1a_2a_3a_4) = \frac{u_1}{d} = 61 & \frac{p_4}{q_4} = \frac{72}{61}.
\end{array}
$$

Of course, 72/61 is exactly u/v.

From Theorem 4.4.2, plugging in a_j for X_j, we can say

$$(-1)^{i+1} = Q[0,i]Q[1,i-1] - Q[0,i-1]Q[1,i] = p_iq_{i-1} - p_{i-1}q_i.$$

Example 2.4. Once again using $\frac{216}{183}$,

$$
\begin{array}{rll}
(-1)^{1+1} = & p_1q_0 - p_0q_1 = & 6*1 - 5*1 = 1 \\
(-1)^{2+1} = & p_2q_1 - p_1q_2 = & 7*5 - 6*6 = -1.
\end{array}
$$

This is saying that the cross product of these fractions alternates between 1 and -1. This is the end of the material we cover in Chapter 4.