2.4 Linear Congruential Generators with Known Module

This section uses elementary methods only and is independent of the general theory from the preceding sections of Chapter 2.

Assume the parameters a and b of the linear congruential generator $x_n = ax_{n-1} + b \bmod m$ are unknown, whereas the module m is known.

We'll show that for predicting the complete output sequence we only need 3 successive elements x_0, x_1, x_2 of the sequence, even for a composite module m. Starting with the relation

$$x_2 - x_1 \equiv a(x_1 - x_0) \pmod{m}$$

we immediately get (assuming for the moment that $x_1 - x_0$ and m are coprime)

$$a = \frac{x_2 - x_1}{x_1 - x_0} \bmod m,$$

where the division is $\bmod m$ (using the extended Euclidean algorithm). The increment b is given by

$$b = x_1 - ax_0 \bmod m.$$

So we found the defining formula and may predict the complete sequence.

A typical tool for this simple case was the **sequence of differences**

$$y_i = x_i - x_{i-1} \quad \text{for } i \geq 1.$$

It follows the rule

$$y_{i+1} \equiv ay_i \pmod{m}.$$

Note that the y_i may be negative lying between the bounds $-m < y_i < m$. Since m is known we might replace them by $y_i \bmod m$, but this was irrelevant in the example, and for an unknown m—to be considered later on—it is not an option.

Lemma 6 (on the sequence of differences) *Assume the sequence (x_i) is generated by the linear congruential generator with module m, multiplier a, and increment b. Let (y_i) be the sequence of differences, $c = \gcd(m, a)$, and $d = \gcd(m, y_1)$. Then:*

(i) *The following statements are equivalent:*
  - (a) *The sequence (x_i) is constant.*
  - (b) $y_1 = 0$.
  - (c) $y_i = 0$ *for all i.*

(ii) $\gcd(m, y_i) \mid \gcd(m, y_{i+1})$ *for all i.*

(iii) $d \mid y_i$ *for all i.*

(iv) *If* $\gcd(y_1, \ldots, y_t) = 1$ *for some* $t \geq 1$*, then* $d = 1$*.*
(v) $c|y_i$ *for all* $i \geq 2$*.*
(vi) *If* $\gcd(y_2, \ldots, y_t) = 1$ *for some* $t \geq 2$*, then* $c = 1$*.*
(vii) $m|y_i y_{i+2} - y_{i+1}^2$ *for all* i*.*
(viii) *If* $\tilde{a}, \tilde{m}$ *are integers,* $\tilde{m} \geq 1$*, with* $y_i \equiv \tilde{a} y_{i-1} \pmod{\tilde{m}}$ *for* $i = 2, \ldots, r$*, then* $x_i = \tilde{a} x_{i-1} + \tilde{b} \bmod \tilde{m}$ *for all* $i = 1, \ldots, r$ *with* $\tilde{b} = x_1 - \tilde{a} x_0 \bmod \tilde{m}$*.*

Proof. (i) Note that $y_i = 0$ implies that all following elements are 0.

(ii) If e divides y_i and m, then it also divides $y_{i+1} = a y_i + k_i m$.

(iii) is a special case of (ii).

(iv) follows from $d| \gcd(y_1, \ldots, y_t)$, and this, from (iii).

(v) Let $m = c\tilde{m}$ and $a = c\tilde{a}$. Then $y_{i+1} = c\tilde{a} y_i + k_i c\tilde{m}$, hence $c|y_{i+1}$ for $i \geq 1$.

(vi) follows from $c| \gcd(y_2, \ldots, y_t)$ and this, from (v).

(vii) $y_i y_{i+2} - y_{i+1}^2 \equiv a^2 y_i - a^2 y_i \pmod{m}$.

(viii) by induction: For $i = 1$ the assertion is the definition of $\tilde{b}$. For $i \geq 2$ we have

$$x_i - \tilde{a} x_{i-1} - \tilde{b} \equiv x_i - \tilde{a} x_{i-1} - x_{i-1} + \tilde{a} x_{i-2} \equiv y_i - \tilde{a} y_{i-1} \equiv 0 \pmod{\tilde{m}},$$

as claimed. $\diamond$

The trivial case of a constant sequence merits no further care. However it shows that in general the parameters of a linear congruential generator are not uniquely determined by the output sequence. For the constant sequence may be generated with an arbitrary module m and an arbitrary multiplier a if only the increment is set to $b = -(a-1)x_0 \bmod m$. Even if m is fixed a is not uniquely determined, not even $a \bmod m$.

Previously we considered the case where y_1 and m are coprime, yielding $a = y_2/y_1 \bmod m$. In the general case it might happen that division $\bmod m$ is not unique. This happens if and only if m and y_1 have a non-trivial common divisor, hence $d = \gcd(m, y_1) > 1$. The **sequence of reduced differences** $\bar{y}_i = y_i/d$ (see (iii) in Lemma 6) then follows the recursive formula

$$\bar{y}_{i+1} \equiv \bar{a} \bar{y}_i \pmod{\bar{m}}$$

with the reduced module $\bar{m} = m/d$ and reduced multiplier $\bar{a} = a \bmod \bar{m}$, from which we get a unique $\bar{a} = \bar{y}_2/\bar{y}_1$. Setting $\tilde{a} = \bar{a} + k\bar{m}$ with an arbitrary integer k and $\tilde{b} = x_1 - \tilde{a} x_0 \bmod m$, from Lemma 6 (viii) we also get $x_i = \tilde{a} x_{i-1} + \tilde{b} \bmod m$ for all $i \geq 1$. This proves:

Proposition 7 *Assume the sequence* (x_i) *is generated by a linear congruential generator with known module* m*, but unknown multiplier* a *and increment* b*. Then the complete output sequence is predictable from its first three* *elements x_0, x_1, x_2. If the sequence (x_i) is not constant, then the multiplier a is uniquely determined up to a multiple of the reduced module $\bar{m}$.*

Thus also in this situation we sometimes have to content ourselves with predicting the sequence without revealing the parameters used for its generation. Here is a simple concrete example: For $m = 24$, $a = 2k + 1$ with $k \in [0 \dots 11]$, $b = 12 - 2k \bmod 24$, and initial value $x_0 = 1$ we always get the sequence $(1, 13, 1, 13, \dots)$.